

THE
CONTINUUM
COMPANION TO
PHONOLOGY

EDITED BY
NANCY C. KULA , BERT BOTMA
AND KUNIYA NASUKAWA

Continuum Companion to Phonology

Also in the Continuum Companions series

Continuum Companion to Discourse Analysis

Edited by Ken Hyland and Brian Paltridge

Continuum Companion to Historical Linguistics

Edited by Silvia Luraghi and Vit Bubenik

Continuum Companion to Phonetics

Edited by Rachael-Anne Knight, Patricia Ashby and Mark J. Jones

Continuum Companion to Phonology

Edited by Nancy C. Kula, Bert Botma and Kuniya Nasukawa

Continuum Companion to Second Language Acquisition

Edited by Ernesto Macaro

Continuum Companion to Syntax and Syntactic Theory

Edited by Claudia Parodi and Silvia Luraghi

Continuum Companion to TESOL

Edited by Jun Liu

Continuum Companion to Translation Studies

Edited by Jody Byrne and John Kearns

Continuum Companion to Phonology

Continuum Companions

Edited by

Nancy C. Kula, Bert Botma and
Kuniya Nasukawa



continuum

Continuum International Publishing Group

The Tower Building 80 Maiden Lane
11 York Road Suite 704
London SE1 7NX New York, NY 10038

www.continuumbooks.com

© Nancy C. Kula, Bert Botma, Kuniya Nasukawa, and contributors 2011

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage or retrieval system, without prior permission in writing from the publishers.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

ISBN: 978-0-8264-3423-4 (hardcover)

Library of Congress Cataloging-in-Publication Data

Continuum companion to phonology / edited by Nancy C. Kula, Bert Botma and Kuniya Nasukawa.

p. cm. – (Continuum companions)

Includes bibliographical references and index.

ISBN 978-0-8264-3423-4 (hardcover)

1. Grammar, Comparative and general–Phonology. I. Kula, Nancy Chongo, 1971–
II. Botma, Engbert Doede, 1972– III. Nasukawa, Kuniya, 1967– IV. Title. V. Series.

P217.C668 2010

414–dc22

2010015145

Typeset by Newgen Imaging Systems Pvt Ltd, Chennai, India

Printed and bound in Great Britain by CPI Antony Rowe, Chippenham, Wiltshire

Contents

Notes on Contributors	vii
Editors' Introduction	xi

Part I: Research Methodology

1 Methodology in Field Phonology	3
<i>Mary Pearce</i>	
2 Methodology in Phonological Acquisition	16
<i>Tania S. Zamuner and Elizabeth K. Johnson</i>	

Part II: Research Issues

3 Features	33
<i>Bert Botma, Nancy C. Kula and Kuniya Nasukawa</i>	
4 Syllables	64
<i>Péter Szigetvári</i>	
5 Stress	95
<i>Diana Apoussidou</i>	
6 Derivations	135
<i>Patrik Bye</i>	
7 Constraint-Based Phonology	174
<i>Christian Uffmann</i>	
8 Phonetics-Phonology Interface	202
<i>Silke Hamann</i>	
9 Syntax-Phonology Interface	225
<i>Anthi Revithiadou and Vassilios Spyropoulos</i>	

Contents

10 Sign Language Phonology <i>Chloë R. Marshall</i>	254
11 Phonological Acquisition <i>Nivedita Mani</i>	278
12 Second Language Phonology <i>Heidi Altmann and Barış Kabak</i>	298
13 Phonological Disorders <i>Dirk-Bart den Ouden</i>	320
Part III: New Directions	
14 Laboratory Phonology <i>Taehong Cho</i>	343
15 Usage-Based Phonology <i>Daniel Silverman</i>	369
Part IV: History of the Field	
16 Issues in the Development of Generative Phonology <i>Tobias Scheer</i>	397
Bibliography	447
Language Index	513
Subject Index	515

Notes on Contributors

Heidi Altmann is a lecturer of English and General Linguistics at the University of Konstanz, where she has been teaching phonology, second language acquisition and sociolinguistics. Her research interests comprise of second language prosody, phonology-morphology interface, and language and gender.

Diana Apoussidou works as a postdoctoral researcher at Utrecht University. Her main research interest is the learnability of phonological structure and underlying representations, with a focus on prosody and metrical properties of language. She is currently working on a model of acquiring a mental lexicon as part of a 5-year research project on phonotactic constraints in speech segmentation funded by the Netherlands Organisation for Scientific Research (NWO).

Bert Botma is a lecturer of English language and linguistics at the University of Leiden, where he teaches courses in phonology, phonetics, morphology and historical linguistics. His main research interest is in segmental phonology. He is currently conducting research on the phonological and phonetic properties of sonorants, under the auspices of the Netherlands Organization for Scientific Research.

Patrik Bye is a researcher at the University of Tromsø and Center for Advanced Study in Theoretical Linguistics (CASTL) in Northern Norway. Bye's research has focused on syllable structure and quantity in Finnic and Saami, accent in North Germanic, phonological opacity, and the interaction of phonology with morphology, lexicon and syntax. He is currently involved in a project investigating the intonation of the Saami languages.

Taehong Cho is associate professor in the Department of English Language and Literature, and the director of the Phonetics and Psycholinguistics Lab at Hanyang University, Seoul, Korea. He earned his Ph.D. degree in phonetics at the University of California at Los Angeles (UCLA) and worked at the Max Planck Institute for Psycholinguistics. His main research interest is in the

interplay between prosody, phonology and phonetics in speech production and its perceptual effects in speech comprehension. He is currently teaching courses in phonetics, laboratory phonology and psycholinguistics.

Silke Hamann is junior professor of phonology and phonetics at the Heinrich-Heine University Düsseldorf. Her research is on the phonetics-phonology interface, especially on the perception of speech sound and on diachronic changes in the mapping between auditory cues and segmental representations. She teaches courses in phonetics, phonology and computational linguistics. She is co-editor of *Phonology in Perception* (Mouton de Gruyter 2009).

Elizabeth K. Johnson is a professor of Psychology at the Mississauga campus of the University of Toronto. Her main interests are in the area of early language acquisition and developmental speech perception. Johnson's Infant Language and Speech Laboratory is funded by the Natural Sciences and Engineering Research Council and the Canada Foundation for Innovation.

Bariş Kabak is assistant professor of English and General Linguistics at the University of Konstanz. His areas of research are phonology and its interfaces with phonetics and morphosyntax, second language speech and laboratory phonology. Some of his current research includes second language speech perception and production, and psycholinguistic aspects of prosodic patterns and their subregularities.

Nancy C. Kula is a lecturer in phonology at the department of Language and Linguistics of the University of Essex. Her research interests are theoretically in representation-based approaches to phonology and empirically on Bantu languages. She is interested in segmental phonology, the phonology-morphology interface and the role of prosody in the phonology-syntax interface. She is currently co-authoring *Key Terms in Phonology* with Continuum.

Nivedita Mani is a junior professor in Language Acquisition at the University of Göttingen. Her main research interests centre on investigating infants' attention to the phonological structure of words in their lexicons. She is currently working on using priming paradigms to examine the formation of phonological neighbourhoods in infant and adult lexicons with eye tracking and EEG data.

Chloë R. Marshall lectures in developmental psychology and language acquisition at City University London. Her research focuses on the interactions between phonology and morphology during development, and on the relationship between phonological working memory and language acquisition. She studies these issues in children with language impairments and in deaf children who use sign language.

Kuniya Nasukawa (Ph.D. in Linguistics, University College London, University of London) is professor of linguistics at Tohoku Gakuin University, Sendai, Japan. He specializes in research on phonological theory, especially the interaction between prosody and melody. He is the author of *A Unified Approach to Nasality and Voicing* (Mouton de Gruyter 2005) and co-editor of *Strength Relations in Phonology* (Mouton de Gruyter 2009).

Dirk-Bart den Ouden is assistant professor at the University of South Carolina, though at the time of writing he was a research associate at Northwestern University. He conducted research on phonological representations in child language acquisition and creolization, before specializing in phonological and syntactic disorders in aphasia. He now uses behavioral as well as functional imaging methods to investigate the neurophysiological substrates of language production and processing, in aphasic speakers and healthy adults.

Mary Pearce works as a linguistic consultant with SIL in Chad. She is an honorary research associate of University College London where she completed her Ph.D. in 2007. She teaches phonology at the Université de N'Djaména, Chad and as part of SIL programmes and workshops in Africa and the United Kingdom. Her main research topics include phonology, tone, vowel harmony and reduction, metrical structure, acoustic phonetics, sociolinguistics, Kera (spoken in Chad), Chadic languages and African languages. Her field experience includes 10 years living in a village location among the Kera.

Anthi Revithiadou, currently assistant professor at the Aristotle University of Thessaloniki, received her Ph.D. from Leiden University/HIL in 1999. Her main research interest is in the morphosyntax-phonology interface. She is pursuing her work in the theoretical context of Optimality Theory, with emphasis on the structure of phonological representations and the issue of parallel grammars in contact-induced systems and in acquisition. She has published in journals and edited volumes. She teaches courses in phonology, morphology and typology.

Tobias Scheer is currently Directeur de Recherche at the CNRS in France. He works at the laboratory Bases, Corpus, Langage (UMR 6039) in Nice, where he is the director. Being a phonologist, his main interests thematically lie in syllable structure and in the (Western) Slavic family in terms of language scope. He is committed to Government Phonology (GP) and hence to the idea that representations have an independent and unoutrankable role in phonology. In 2004, he published a monograph on CVCV phonology (a particular development of GP) with Mouton de Gruyter, and is currently preparing the second volume devoted to the interface of phonology and morpho-syntax.

Daniel Silverman teaches phonology and phonetics at San Jose State University. He has published widely on phonology, especially focusing on the diachronic interaction of phonetic and functional pressures on linguistic sound systems. He is the author of *A Critical Introduction to Phonology: of Sound, Mind, and Body* (Continuum 2006), and is currently writing a book on Neutralization (Rhyme and Reason in Phonology) (Cambridge University Press).

Vassilios Spyropoulos is lecturer of generative syntax at the University of Athens. His main research interest is in syntactic theory and analysis and the interface of syntax with phonology and morphology. He is currently working on the role of the syntax-phonology interface in the linearization of chains, feature decomposition of case, and raising and control phenomena in Greek. He teaches courses on syntax, morphology and grammatical description of Greek.

Péter Szigetvári is associate professor at Eötvös Loránd University, Budapest. His interests include consonant lenition, markedness and phonotactics. He lectures on phonology, IT and typography.

Christian Uffmann is lecturer in linguistics and English Language at the University of Sussex. His main research interests are in phonological theory, especially in the role of representations in constraint-based phonology, and in English phonology, especially in contact phenomena, including creole formation and loanword adaptation. More recently, he has also begun to look at sociolinguistic variation and change.

Tania S. Zamuner is assistant professor at the department of linguistics of the University of Ottawa. Her areas of interest include language acquisition, developmental speech perception and production, psycholinguistics and lexical development. Her current research focuses on phonotactic development in speech perception and production, the role of production in word learning and cross-linguistic lexical development in typical children and in children with speech delay. She is author of *Input-based Phonological Acquisition* (Routledge 2003).

Editors' Introduction

When we first embarked on this project it was a fairly easy task to decide what topics should be included in *Continuum Companion to Phonology*, but the choice of what approach to these topics to adopt was rather more complicated. We contemplated making the companion theoretically focused but in the end we settled for a more neutral approach that allowed for the discussion of many different approaches to phonology. This companion therefore rather than generally advocate a particular approach to phonology instead aims to highlight areas of interest and importance in phonological research while also showcasing the different approaches that may be adopted to these issues. The choice of authors is therefore a reflection of this diversity and we hope that you will be inspired and challenged by some of the ideas and perspectives presented.

The volume is divided into four parts that present a holistic approach to phonology covering both research and methodological issues. The following provides a synopsis of each part of the companion and the chapters contained therein.

Part I deals with methodology in phonological research and offers two papers on field phonology and experimental work in phonological acquisition. **Pearce** (Chapter 1) provides guidance and insight on issues that must be considered if one embarks on collecting phonological data in a particular speech community. The chapter, which is informed by her own experiences, also offers practical examples in the elicitation of segments, syllable structure and tone, highlighting the pitfalls that must be avoided in fieldwork. **Zamuner and Johnson** (Chapter 2) review some experimental techniques for probing the development of speech perception and production in children from infancy to the age of 2. The chapter explores the relationship between speech perception and production, focusing on the Anticipatory Eye Movement Paradigm (for investigating perception) and Word Elicitation and Non-Word Repetition Tasks (for examining production). The chapter offers useful methodologies for researchers who are interested in these intertwined areas of early phonological acquisition.

Part II deals with the core issues that continue to form the basis of current phonological enquiry. There are 11 chapters that cover the main areas of phonological research; the first five chapters deal with issues internal to phonology as a system covering features, syllables, stress, derivations and constraints. The next six chapters deal with phonology in a wider perspective and in particular how it interacts with different language components and issues including phonetics, syntax, first and second language acquisition, sign language and language disorders. The chapters in this part are organized to offer a general introduction to the topic in the initial part and a case study of a particular phenomenon in the latter part. We aimed in this way to provide not only an overview but also a concrete illustration of the particular perspective that is fostered.

Botma, Kula and Nasukawa (Chapter 3) on phonological primes defend a privative view of phonological features showing that quite complex interactions between voicing and nasality that are usually argued to support binary features can be accounted for under a privative view that assumes enriched sub-segmental representations that are subject to dependency relations. Under the proposed view the same phonological prime is argued to contribute different acoustic information to the eventual signature of a segment by virtue of its position within segmental representation. **Szigetvári** (Chapter 4) presents arguments that illustrate why the notion of a syllable is descriptively useful in the explanation of various cross-linguistic phonological phenomena. He furthermore scrutinizes the formal status of syllables in phonological representation arguing for an alternative model of representation without syllables. In his strict CVCV model the phonological skeleton consists of strictly alternating consonantal and vocalic positions regulated by licensing and government relations. In the final part of the chapter, Szigetvári demonstrates that for a number of recurrent phonological phenomena, the CVCV model offers better insight. **Appousidou** (Chapter 5) provides an introduction to the typology of stress systems and presents arguments for the computational modeling of the acquisition of stress. The larger part of the chapter focuses on a case study of stress patterns in Greek where she attempts to mirror the difficulty and the variation exhibited by Greek speakers in learning shifting stress in an acquisition simulation. The chapter by **Bye** (Chapter 6) is concerned with the nature of phonological derivations and levels of representations. Bye focuses on extrinsic rule interactions and illustrates how these are analysed in rule-based phonology and in Optimality Theory. Bye provides substantial discussion of other principles of ordering such as persistent rules, intrinsic ordering and cyclicity in a quest to answer the fundamental question of whether linguistically significant generalizations should be stated on levels or via derivations. **Uffmann** (Chapter 7) in his chapter on constraints discusses the gradual rise of constraints in generative phonology, from their use as morpheme structure

conditions to their central role in frameworks such as Declarative Phonology, Government Phonology and, most notably, Optimality Theory. In the final part of his chapter Uffmann provides an overview of the main tenets of Optimality Theory, in which he focuses on such aspects as the evaluation and violation of constraints and the notions of faithfulness, markedness and Richness of the Base.

The next six chapters on linguistic interfaces start with **Hamann** (Chapter 8) who presents an overview of issues in the phonetics-phonology interface with the conclusion that the characterization of how phonetics interacts with phonology is heavily dependent on the phonological assumptions that are made. In the final part of her chapter, Hamann advocates the need for explicit models and formalizations of the interface and offers a detailed description of BiPhon, a modular approach to phonetics and phonology that incorporates both production and comprehension. **Revithiadou and Spyropoulos** (Chapter 9) look at the interaction of phonology and syntax and elaborate on ideas developed in prosodic phonology and associated phonological phrasing defining the interface between the two modules. They provide an overview of the main approaches to the interface; direct access and the prosodic structure hypothesis, including end-based and edge-based mapping. In the second part of the chapter they discuss a case study on the variant positional interpretation of Greek subjects as following from the interaction of prosodic constraints within the copy theory of movement and the multiple spell-out hypothesis within Chomsky's Minimalist Syntax. **Marshall** (Chapter 10) offers a lucid discussion of some of the phonological properties of sign language. She focuses on the similarities and contrasts between spoken and signed modalities of language, observing among other things that sign languages involve a higher degree of simultaneous (or non-linear) processing and iconicity than spoken ones. Marshall concludes her chapter by describing the preliminary results of a case study on lexical segmentation in British Sign Language (BSL). The results of this study suggest that signers are faster and more accurate in identifying signs that are preceded by non-existing but possible signs, similar to what has been observed in spoken language. **Mani** (Chapter 11) presents a comprehensive view of first language acquisition of phonology and provides a well-rounded overview of the issues that are central to phonological acquisition. She draws on a wide range of experimental work demonstrating that a number of issues including acoustic information, frequency, type of phonemic inventory and timing of exposure, all converge in phonological development. **Altmann and Kabak** (Chapter 12) highlight the many ways in which second language acquisition differs from first language acquisition touching on the issue of whether ultimate attainment is possible in second language. They offer an overview of second language research and the acquisition of segmental contrasts, syllable structure and prosodic structure, demonstrating that perception

plays a greater role in second language phonology than has been assumed in prior research. They discuss issues in the L2 acquisition of stress and the role that the L1 has to play and show from various studies that the perception of stress may in part be influenced by the kind of stress system in the L1. **Den Ouden** (Chapter 13) provides an overview of research on phonological impairments showing that this research is characterized by two perspectives: research which uses phonological and psycholinguistic theories to inform the analysis of impaired processing, and research which uses data from impaired processing to inform phonological and psycholinguistic theorizing. Den Ouden investigates the impact of aphasia on syllable structure, discussing two case studies involving Dutch aphasics which suggest that for these speakers hierarchical syllable structure comes into play only post-lexically, at the level where segments are mapped onto syllable slots.

Part III of the companion offers new directions in phonological research that reflect the way in which the field is expanding. We focus here on two main areas. There has been a steady increased interest in experimental work in phonology in recent years reflected in work in laboratory phonology, on the one hand, and also a resurgence of more exemplar/usage-based approaches to linguistics in general leading to usage-based approaches to phonology. **Cho** (Chapter 14) provides a brief history of the field of Laboratory Phonology noting that this approach has been motivated by the increasing awareness that many non-contrastive phonetic events are not automatic or universal, but are instead either systematically linked with phonological contrasts or governed by language-specific phonetic rules. The second part of Cho's chapter is concerned with the interface between prosody and phonetics where he provides a detailed discussion of the manifestation of prosodic structure in speech production. He also considers how aspects of prosodic structure may help listeners in lexical segmentation. **Silverman** (Chapter 15) in the second chapter in this part traces back the usage-based perspective to phonology to such nineteenth-century scholars as Baudouin de Courtenay and Kruszewski. As Silverman shows, the work of these scholars was greatly influenced by Darwin's theory of evolution, in its 'emphasis on slow-going diachronic pressures that may shape and re-shape the linguistic system, due to specific patterns of use and disuse'. Silverman shows that the same emphasis is found in more recent usage-based approaches such as the work of Bybee. In the second part of his chapter, Silverman goes on to discuss the main tenets of the research programmes of two modern progenitors of usage-based phonology, Ohala and Labov.

Part IV of the companion is dedicated to the evolution of phonology as a field with the goal of contextualizing the different strands of phonology that can be observed today. **Scheer** (Chapter 16) provides a thoroughly researched and critical chronological development of the field from the late 1960s to current times. Scheer observes that while early phonological work was mainly

concerned with rules (computation) the surge in different approaches in the 1980s theoretical boom was a reaction to expunge rules from phonology and replace them with representations. The field, however, seems to have come full circle by reverting to computation in the form of constraints with the advent of anti-derivational Optimality Theory. Scheer points out that the tension between computation and representation is one that is bound to remain but that an adequate theory probably needs a measure of both.

As these short summaries attest this companion is filled with many gems that we hope you will enjoy discovering.

We would like to express our sincere thanks and acknowledgements to the many phonologists all around the globe who gracefully accepted our demands on their time to act as reviewers for the chapters that make up this book. Needless to say, their many insightful comments have played a pivotal role in shaping this work into valuable research that we hope will stand the test of time.

We also express our thanks for grants held during the production of this book to the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government (grant numbers 19520429/22320090 awarded to Kuniya Nasukawa) and to the Netherlands Organization of Scientific Research (NWO) (grant for project number 2008/01214/GW awarded to Bert Botma).

Nancy C. Kula, Bert Botma and Kuniya Nasukawa

This page intentionally left blank

Part I

RESEARCH METHODOLOGY

This page intentionally left blank

1 Methodology in Field Phonology

Mary Pearce

Chapter Overview

1. Introduction	3
2. General Considerations for Fieldwork	4
3. Phonological Issues in Fieldwork	8
4. Conclusion	15
5. Acknowledgements	15

1. Introduction

There is a tendency among linguists to polarize towards either a theoretical or descriptive approach and this means that most books on phonology will either discuss theoretical issues with just a minimum number of data examples, or it will give detailed data with only a few references to linguistic theory. But this handbook acknowledges the scope for descriptions with a solid theoretical analysis and for theories which are backed up by solid and convincing data. In order to produce a well-analysed description or a theory based on empirical evidence, it is necessary to carry out field research.

In this brief overview of phonological field methods, I am assuming that the reader is probably a phonologist with a theoretical background who would like to collect data in order to describe some aspect of the phonology of a specific language or to find support for a theory. The data collection could take place in a multicultural city with one or two speakers of the language, or in a speech community, possibly in a rural setting. Some of the ideas presented here do not apply to all situations. Evidently, if an anechoic room is available, some of the comments about recording equipment are not necessary, but I will assume that the linguist may need to take all of their recording equipment with them. As Bowerman (2008b) points out, it is not always necessary to travel to a remote part

of the world to do fieldwork, but measuring just one speaker outside of their normal language context is unlikely to give much insight. Therefore, it is assumed in this chapter that there is a speech community involved and that the linguist may well have travelled to work with them. If the language under consideration is either endangered or otherwise in a contact situation, Bower (2008b) is a helpful resource to refer to. For endangered languages, there is the whole question of what should be documented and described. This is beyond the scope of this short article, but the fieldworker should consider their responsibility towards the documentation and also be aware that they may not be able to return at a later date for more data. They may therefore have to concentrate on recording more dialogues and general data rather than focusing exclusively on one small part of the phonology.

I will begin with an overview of general principles and considerations which apply to all types of fieldwork, and then move on to consider a few areas of phonology where specific guidelines may be of use. There are three books which have been written on this topic, which I find particularly helpful as reference books. These are Ladefoged (2003), Crowley (2007) and Bower (2008a). Chapter 5 of the latter is particularly relevant for Phonology as is the regularly updated website and appendix with a phonological checklist.

Note also that various software options are available which can help considerably in fieldwork. As well as the acoustic analysis tools and data management tools provided by university departments, research institutes and SIL International, there is also an increasing selection of resources being made available from the endangered language centres. These tools help with electronic descriptions, data from related languages, guidelines for research and tools to speed up research.

2. General Considerations for Fieldwork

If the time for fieldwork is limited, some feel that learning the language is an optional extra which they cannot afford to engage in, but it is worth attempting. The speakers appreciate the effort and will probably respond with more enthusiasm to the tasks they are given and in learning the language, the linguist will develop a greater appreciation for the areas where more analysis is needed. Clearly, if the length of fieldwork is less than 6 months, the time is probably better spent in elicitation and analysis, but even in this case, it is still worth learning greetings and simple phrases as this helps in relationship building.

In any case, respect for the speakers of the language is essential. This includes respect for the phonological insights that the speaker might have. It is not normal for a speaker to be able to verbalize all the phonological categories in

their own language, even if linguistically trained, but speakers can often make sound judgements as to whether two things are different from each other, and it is worth investigating these judgements further. The participatory approach as advocated by Kutsch Lojenga (1996) is helpful, where the opinion of the speaker is sought when trying to decide on contrasts. Of course this should not be the only method used to find contrast: British speakers may feel that *cheetah* and *cheater* are pronounced differently and that *those* and *thing* begin with the same sound, having been influenced by the orthography. So investigations of this kind should always include recordings.

Before conducting fieldwork, the researcher should consider what material will be made available to the language community after the research. This might include a lexicon, copies of files, a phonological description etc. Feedback can also be given during the fieldwork. Many speakers are interested in the results of the research and in learning more about their language. In experiments which involve a number of participants, it is common practice to just use initials when referring to subjects, but language consultants should be acknowledged.

For all fieldwork, it is becoming increasingly important to have ethical and government permission for linguistic work. More details on this are usually available from universities or research bodies that support fieldwork. The usual request is that the linguist gets written consent from anyone that they record. I have found that in some locations, having a form to sign can be culturally inappropriate – so my approach has been to record them giving permission (in their own language). In this way I have the permission and no one is upset, and in addition, I have more data in the language which I can analyse. Related to this, there are issues concerning how much to pay people who help with research. For this, it is best to get advice from other linguists who are working in similar settings.

Before embarking on fieldwork, a great deal of planning is necessary. The plans should be flexible as they are likely to change as soon as some data is collected. It is good to plan in time to consider the results, to reanalyse, and to devise new tests during the fieldwork in light of what the early sessions reveal. This means that it is important to have a system for recording field notes involving separate notebooks for transcriptions and glosses, for plans for each session with lists of data to record etc., and for ongoing analysis. In the initial planning, it is helpful to write a list of research questions, and then to decide what data are needed to answer those questions.

The right level of phonetic transcription will probably depend on the research questions. The linguist must decide how much to simply listen out for things that they expect and write a transcription accordingly, and how much to organize the material so that the transcription is clear for inquiries other than the one they are engaged in. However good the phonetician, all transcriptions need to be treated with an open mind as until the phonological analysis is complete,

some details will never be certain and others will only get resolved with the use of speech analysis software. Casali (2000, 2008) notes the need to maintain a healthy distrust of the data and to be willing to revise it.

It is clearly important to collect enough data to support your claims. But there should also be enough data to support theories that may not have developed yet. For this, it can be helpful to have a body of general data to fall back on, such as word lists and natural speech. If possible, these data should be published after the field trip so that other researchers can verify the data and the claims, and so that the local community can benefit. Several word lists are available from research institutions involved in descriptive linguistics. It can also be helpful to consider discussing some semantic domains in more detail to get a longer list. For this the use of books on topics such as plants or birds can be helpful. To vary the activity, I also devise games where speakers draw fun pictures and then describe them or make up stories. If I am focusing on grammatical tone, I include elements which will produce words in that category. If I need nouns with the structure CVCV, for example, I draw a map with various objects beside the road (carefully choosing objects which have names of the target structure). Then I ask for the speaker to tell a story of someone who travels the length of the map passing the objects.

When eliciting phonetics and sometimes phonology, it can be helpful to try out various versions of an utterance to see which is most acceptable. Ladefoged (1997) says 'One of the most efficient procedures for getting results in the field is to test different hypotheses by trying out various vocal gestures of our own.' This can of course be overdone if we then don't listen to the proper pronunciation adequately, but as long as the speaker is happy to tell us when we are wrong or to choose between two options, this can be effective.

2.1 Acoustic Phonetics and Recordings

Although I discuss this only briefly, the use of instrumental analysis in research is vital. This has been noted as far back as Pike, Ralph and Bascum (1959). Baart (2001) claims that acoustic phonetics in fieldwork adds accuracy, precision and objectivity. It is useful for ascertaining pitch height and movements, duration and quality of vowels and consonants, measurements of Voice Onset Time (VOT) and syllable length, aspiration, voice quality and phonation.

While acoustic measurements are very useful, the researcher should guard against such assumptions as 'the pitch extraction will reveal the correct phonetic tones', or 'the formants will tell me which vowel it is'. There are two dangers here: first, the fieldworker can rely too much on pitch tracks rather than their own hearing, and secondly, they can confuse phonetic data with

phonological analysis. But once phonological hypotheses have been developed, the acoustic data can provide useful support for the claims. For example, if the claim is that there are six contrastive vowels, you would expect to find clustering of the formants around these positions – but it is not valid to find six clusters and therefore assume that there are six vowels.

The field linguist must establish what level of quality to aim for in a recording. It is clear that we would all like the highest level possible, but this has to be weighed up against ambient noises such as goats and traffic and the fact that if recordings are done outside (where there is enough light), the microphone may pick up the sound of the wind or neighbours etc. There is also the issue of how intrusive the recording equipment looks. I have at times recorded a group with a digital recorder with internal microphone because that appears to be just a cassette recorder. I may have lost something in quality, but the speakers felt more at ease and they were able to talk naturally. It may be that a mixture of approaches is called for with different groups and different research questions. It is also a good idea to have a second recorder as backup in case the first fails for some reason.

The recordings should be digital and the recording equipment should be user friendly to avoid disappointment (which means avoiding small recording machines with several commands on one button). Always test the recordings for ambient noise such as light fixtures which may affect the pitch track. Recordings in WAV format are normally better than MP3 recordings if you are carrying your own data on a computer, but if data is being sent to other people electronically, it may be that WAV files are too large, in which case MP3 recordings (although they are compressed) can be used for most purposes. Sometimes MP3 is better because the compression loses some of the extraneous noise. This will depend on the purpose of the recording.

There are several data recorders and dictaphones on the market now, some of which are better than others, so it is good to ask for advice from field linguists before purchasing one. Elements to consider are: the power source, for example, AA batteries if in a remote place, USB download, WAV files, manual recording levels – an indication of overload in volume, 44 kHz or higher, good microphone, easy controls and memory for previous settings, some sort of playback. Some portable computers are also fine for recording with an external microphone, but others add extra noise so this must be tested in advance. For the analysis, there are various computer programs available which can both record the speech and present the wave forms as spectrograms and pitch tracks. Some offer extra analysis tools and the chance to manipulate the recording for use in perception tests etc. These programs are extremely useful to the field linguist, but it is important not to rely on the visual data alone. So the linguist needs to listen first and then use these programs as a backup. A laryngeograph can be useful if

breathily voice is interacting with tone, and if speakers are happy about using it (and about giving consent in written or recorded form). Most of this equipment takes practice to use, so good preparation time is essential.

All recordings should include the date, the speaker, reference numbers and glosses. It is advisable to speak these data into the microphone at the beginning of every recording session as it makes labelling easier. The files should be kept in well-ordered folders with clear marking of what they contain. In my field trips, I have recorded about 1 megabyte of data per week. This amount of material requires good organization.

It may be that the research includes some statistical analysis. Some of this can be done while still in the field situation, but it is more likely that the majority will be done at a later date. Therefore, it is essential to make sure that the details of the speakers are recorded, such as age, gender, education, dialect, parentage etc., as it is otherwise hard to find these facts at a later date. Special care should be taken if there is unlikely to be an opportunity to return to the same place. Similarly, care should be taken if the language is endangered and further research may be impossible.

If the language is endangered, it is worth considering carefully what material should be documented, even if this is not directly linked to the research question. Endangered language courses and departments can provide details on what information is most important, and a researcher going to work on an endangered language should really enrol on at least a short course in documentation before doing fieldwork. If this is not possible, the relevant departments can still provide helpful suggestions and they may know other linguists in the area.

3. Phonological Issues in Fieldwork

3.1 Contrast and Variation

In most theoretical work, data sets are presented as clean, ordered sets so that the contrasts and variation are easy to spot. In reality, this task can be quite complicated. In order to establish the phonemic status of segments, for example, one can create a large chart with a column for each phone and rows for each environment. The chart can then be used by the linguist to indicate minimal or analogous pairs and other items to show that the sound exists in each environment. The cells that remain empty may reveal phones that are in complementary distribution. It is also often helpful to list all the phones that appear on either side of two phones that are likely to be in complementary distribution. The use of a chart thus makes the conditioning environments more apparent than simply looking at all the words in a word list.

For analysing syllable structure you can circle all ambivalent items in a word list, such as vowels that could be glides. Then the analysis of syllable shapes can be based on words with no ambivalent segments.

For comparing contrastive segments or tones, it is helpful to develop some frames to place them in so that the rest of the utterance around the words being considered is always the same. Think of several frames in advance as some frames may not work well for various reasons. It could be that they perturb the patterns too much (particularly for tones), or it could be that the frame seems so unnatural that the speaker cannot say them naturally. I found this with frames such as 'Say . . . again' which produced a major pause before and after the item.

There are a number of games which can be used for phonological purposes. Look first at word games that are already in use in the community. Games such as Pig Latin can reveal helpful information about the phonology of the language. It is also possible to experiment with nonce words to reveal phonological structures. Yip (2003) discusses various ideas along these lines. Loan words (Yip, 2006) can also be useful. They should not be treated as if they are normal words in the language, but if their source and original form is known, the changes can be informative. For example, I noticed in one field trip that words borrowed into Kera from French had a final high tone although this is not the normal pattern in Kera or French. There was also a strong voicing/tone interaction in loans. This told me something about Kera phonological rules and also about the Kera perception of French as having a stressed final syllable (perceived and realized as a high tone in Kera).

3.2 Tone

With all topics, but especially with tone, it is worth researching neighbouring and related languages before beginning work on tone. However, it is also worth keeping an open mind as each language is different and write-ups on tone can differ in their accuracy. The speakers of the language will probably know who is good at producing tone and who would make a good recording. Generally, some are better than others. People who have travelled to cities may lose their tone first before vocabulary is lost if in contact with a non-tonal language.

For musically inclined linguists, it is important to remember that a semitone difference may not be a difference in tone. Linguists who are not musical may fear that they will never hear tone, but a lot can be achieved by asking the speaker for their opinion and in comparing two melodies to decide if they are the same or different. Tone is relative and so must be seen in context. Two words from different recordings should not be compared directly for tone. The linguist

needs to be aware of possible effects from peak delay, interpolation, downstep, intonation, speech styles, speed and excitement. All of these can be tested, but it is best to avoid as many as possible at the beginning. Short recordings are more useful for analysis, but a story is best recorded in one go. Repeated recordings eliminate fluctuations.

For tone analysis, word elicitation is essential. It is best to put all nouns and verbs in a database. This should include the fields of the phonological form, the phonetic form, the surface tone, the underlying tone, the category of speech, the class, the gloss, the CV pattern of the stem and the CV pattern of the word, plus voicing if relevant. In general, it is best to work on two or three syllable nouns first and then verbs, sorting the words into those with a similar structure. The database needs at least 1500 words to be useful for tone analysis. If the nouns have obligatory class markers, these will of course be included in the basic forms. Although morphology is best avoided at early stages in the analysis, there is no point in trying to elicit roots without obligatory affixes. However, if there is a class with a zero morpheme, this can be a useful starting point.

A good system of annotation is essential for tone. It may be helpful, at least at first, to write two parallel lines above each word and then use horizontal dashes at six or seven levels to indicate pitch. This is less misleading than moving straight away into H, M, L notation before analysis.

3.2.1 Tone Frames and Identification of Tones

Tone frames such as the following are useful: Say . . . again, I saw a . . . yesterday, I put a . . . on the table, I am . . . ing now. I was . . . ing yesterday. Two-sided tone frames should have consistent tones on both sides of the word. It is also possible to make good use of one sided frames. Sometimes these perturb the tones less than two-sided frames. Keith Snider (pers. comm.) makes use of the following frames: . . . HH, . . . LL, HH . . . , LL If list intonation or downstep is likely to be a problem, then the words should be tested in different places in the utterance. A long list of words is needed in each frame to be sure that the pattern is a stable one, and also in order to be prepared for one list splitting into two or three when put in a different frame. Clearly, each frame should include words from the same syntactic category as the syntax may introduce extra tones which affect the words under consideration. So a noun followed by an adjective should not at first be compared with a noun followed by a verb.

In some languages it can be easier to start without a tone frame. Although this sounds like a bad idea because tones are relative, and each word may start with a new frame of reference, if the speaker can say several words of a similar structure in a rhythm, any words with a different melody stand out. However, avoid repetition, as the speaker may then form artificial phrases with a phrase intonation. If the linguist is relying on speaker judgements to decide on which tones are which, it can be helpful to write the words on pieces of paper and then

arrange them in piles of words that sound the same or different. This avoids the pitfall of labelling tones too quickly and it identifies the word melodies, which in most languages is more useful than simply identifying the tones. At the end of this process, each group should be retested together to check for any that are in the wrong pile. Once the word melodies have been separated out, they can normally be placed in a relatively symmetrical pattern such as LH HL HH LL. If there is no LH but there is a LM, this is the point at which to decide if a M tone can be explained away. It may be that certain segments such as nasals or voiced obstruents are affecting the melody. This should be clear from a comparison of the words in each pile. This approach also reveals where contours are possible, and it gives access to the information needed to decide on what the tone bearing unit is.

Whistling the melody can be useful if the culture allows it. Note that if the speaker whistles, it may not be the same as what he says, but it can still give useful information, for example, he may whistle the number of syllables in the underlying form. Sometimes the speaker can whistle a fall which he doesn't pronounce. This can give further clues of the underlying pattern. Sophie Salfner (pers. comm.) found this approach particularly useful and relates that it also appealed to children who began to constantly hum tones as a game thereby providing naturally occurring data that could be used for corroboration. If speakers can't whistle or are embarrassed to do so, they may still be willing to hum the melody.

Certain information may be available in an impressionistic way from texts. For example, intonation patterns, downstep, stress, duration features, the effects of speech rate and discourse structure may be clear from the text. But for detailed analysis of tone, it is always best to start with elicited words in isolation or in frames.

In languages where there is a link between tone and other features such as voicing or stress, it is possible to use nonce words to investigate the tonal processes. For example, Kera tone melodies on three syllable words have tone patterns based on the foot structure and voicing. I created nonce words such as <garnatam> and <barantam> which would suggest a L first tone and H last tone based on the voicing of the onset obstruents. I then investigated the middle tone (which always matched the tone of the other syllable in the iambic foot regardless of whether the disyllabic foot was first or last). I also played manipulated recordings of strings of <bapabapaba> type words with the pitch of the <pa> syllables either matching the <ba> syllables or produced about 1 semitone higher. They repeated back to me what they heard. The result was that either the change in pitch and voicing was removed and all syllables were pronounced as <ba> on a level pitch, or the pitch interval was increased for the <pa> syllables. Although experiments of this kind cannot be used alone to test for tonal melodies and processes, they can still provide useful corroborating evidence.

3.3 Segments and Duration

Before investigating the nature of consonants and vowels or syllable shapes, it is worth looking at the literature on related languages to establish what the issues are. But as with tone, an open mind is still important.

Digital recordings can be used to investigate vowel quality, reduction, contrasts, vowel harmony, VOT and formant transitions. They can also be used to compare the durations of epenthetic vowels, transitions, contrastive vowel length, consonants and syllables. The measurements need to be accompanied by well-grounded analysis. For example, in Kera, the transitions (which are not phonological vowels, just consonant releases when moving to a different consonant) can have up to 20 milliseconds duration depending on the surrounding consonants. Epenthetic vowels can have 20–60 milliseconds duration, short vowels in non-head syllables in a foot can have 30–50 milliseconds duration, non-footed vowels have about 50 milliseconds duration and short head vowels can have 50–80 milliseconds duration. Phonologically long vowels have 60–150 milliseconds duration. Given the range of all of these, it is no easy task to sort out the phonetic differences, the effects of stress and the phonological contrasts in the duration. It often requires a lot of data and careful measurement together with an understanding of the phonetic influences on each of these vowel types.

When measuring vowels and tones, it is important to have a consistent method of measurement. For many languages, the best place to measure is past the half way mark for the vowel, but not right at the end. If the formants or pitch track are clearly peaking at another place, it may be necessary to revise the place of measurement, but consistency is important.

The researcher should remember that acoustic measurements cannot always produce an accurate phonological distinction in vowel length. Even in languages with vowel length, some long vowels in fast speech can be shorter than a phonological short vowel in slow speech. Phonological length (monomoraic or bimoraic) must be decided on the basis of analysis which takes into account the tone patterns, permissible structures etc., as well as the duration. Other languages show a great variation in duration due to stress or position in the phrase, but phonologically there is no distinction.

3.4 Foot Structure and Rhythm

In some languages the foot structure or rhythm is clear. In others, there is room for investigation; first, to sort out what the stress patterns are and how feet are formed, and secondly, to note the cues for stress. Stress is usually not a concept that is recognized by the speaker, but it may none-the-less be perceived. I have

worked with speakers who could identify non-footed vowels as being different from other vowels with the same quality. I have also asked people to tap out rhythms and found that they were tapping only heads of feet. However, it is important to experiment with the approaches towards investigating stress. I obtained recordings of phrases said at a slow rate in the hope that this would reveal vowel quality changes, but the speaker said words at a normal rate with pauses in between. I was about to delete the recording when I realized that the pauses were actually between feet rather than between syllables or words. The moral that I took from this is first, that I should not delete any data as it may be useful later, and secondly, that the speaker may be aware of stress at some level without realizing it. It is good to remember that data recordings can be used for several purposes, not just to answer the research question that it was originally recorded for.

3.5 Experiments and Perception

Psycholinguistic and perception tests can be carried out in remote areas, but there are certain challenges that must be faced which are not a problem in a university linguistics laboratory. In remote areas, the power supply may be limited, so it may be necessary to run several people through a test at the same time. When my computer had 2 hours of battery only, I had to ask several people to complete a perception test together in one room. This of course is far from ideal, but sometimes it is that or nothing. My results were still good. Note that group recordings are good for interaction, interest and naturalness, but there is a risk that the recording will be less good because of the distance from the microphone and that each will copy the pronunciation of the first person to speak. The answer to this is recording each person for several phrases at a time (and note which person is speaking).

It can be hard to get the right groups for experiments so that the results are controlled for age, education, gender, dialect and variation among subjects. (See Pearce, 2007, for an example of how these factors can all affect the results of phonological experiments.) For the average phonologist who may not have studied sociolinguistic methods in depth, it is important to research the key issues for such tests before they are carried out, together with the amount of data which will be required to get statistically significant results. As already stated, the biographical data on participants should be collected at the time of doing the test.

In the case of dialect change, it may be that tests could be repeated a number of years later. If this is a possibility, it is important to write up all the details of the test clearly so that a comparative test has some meaning even if carried out by a different linguist. In any case, the conditions of the test including details on

recording equipment etc., need to be described in detail so that the results can be evaluated by other linguists.

3.6 Areal Features

Descriptive linguists are usually well aware of the influences from diachronic processes and language contact. The temptation is sometimes to describe every language in an area in the same way. On the other hand, theoretical linguists can be too ready to find completely new phenomena based on their view of typology without considering the genetic and contact influences which give a justification for what is happening. So it is important to know the characteristics of the language family of the language, and to consider the languages that are in contact and what their characteristics are. It would be helpful too to consider sociolinguistic issues about which languages are likely to influence the community most.

The effects of language contact should be considered in any speaker who is bilingual even if the language being researched is his/her dominant language. It may not affect the phonological grammar, but the phonetics will be affected. For example, tone language speakers can quickly lose their tones when they move to a town where a non-tonal language is spoken. They may retain enough tone to be understood by other speakers (particularly for minimal pairs), but their perception and production of tone may be different from that of a speaker from a rural area. Vowel formants can fluctuate similarly.

In Africa at least, there are several phenomena which seem to take on an areal distribution (Clements and Rialland, 2007), rather than a genetic (Hyman, 1999) or typological (Hyman, 2002) one. One clear example of this is the use of the labio-dental flap (recently introduced into the IPA (International Phonetic Alphabet)) which is found in around 60 languages in Africa from several different language families. The distribution seems to follow the Bantu migration, although many of the languages that make use of the sound are not Bantu (Olson and Hajek, 2004).

A good description needs to distinguish patterns that are routine and expected from those that are surprising and interesting. This applies whether writing a complete phonological grammar or focussing on one aspect in support of a theory. It is therefore necessary to do research before the fieldwork to ascertain what is routine and expected. The fieldworker should not assume that the expected pattern is what they will find, but it is a good starting point. They should test for the expected patterns to see which are attested and then look for unexpected patterns that make this language different from neighbouring languages.

4. Conclusion

Bowern (2008b) notes that the best approach is to have a diverse analysis toolkit with a variety of techniques coupled with good quality recordings and annotation. The techniques should include both observation, experimental and elicited data from a sample of the population. In fieldwork situations, the unexpected is to be expected and you need to be prepared to laugh at yourself. It is also of great importance that the phonologist is both flexible and prepared to abandon one approach in favour of another one. It is also vital that evaluation and careful filing takes place during the fieldwork as data which was collected for one purpose may well lead to new conclusions in another area of research. In most fieldwork situations, we cannot hope for 'perfect' conditions with total silence for recordings and people lining up to do repetitive experiments in the minimum amount of time. But with the downside of fieldwork, there is also the gain of being able to work within a community other than your own and to grow in understanding of not only their phonology but also their way of life. Fieldwork can be very enriching in many ways, and certainly the understanding of phonological systems benefits greatly from linguists who make the effort to do thorough fieldwork.

5. Acknowledgements

I would like to thank several colleagues who gave me suggestions on this topic. They include Mike Cahill, Eric Carlson, Rod Casali, Connie Kutsch Lojenga, Russell Norton, Sophie Salfner, Keith Snider, Paul Unger and Moira Yip.

2 Methodology in Phonological Acquisition

*Tania S. Zamuner and
Elizabeth K. Johnson*

Chapter Overview

1. Introduction	16
2. Methods to Assess Developmental Speech Perception	19
3. Methods to Assess Developmental Speech Production	22
4. Conclusion	28
5. Notes	29

1. Introduction

The emergence of speech perception and language production is integrally related, with the development of the perceptual system in many ways setting the stage for the development of the production system. Surprisingly, despite the necessary link between these two systems, they have traditionally been studied independently of one another. The lack of work focused on understanding how language perception and language production develop in concert is in part due to disciplinary boundaries. Classic work in the area of developmental perception has been most often carried out by psychologists, whereas classic work in the area of developmental production has been most often carried out by linguists. Communicating findings from one field to the other has only recently become commonplace as a growing number of researchers have been trained in both fields. Methodological challenges present another impediment to the integrated study of language perception and production. Many classic methodologies for studying developmental speech perception, on the one hand, are best suited for use with infants, and rely on cross-sectional designs that reveal typical developmental patterns in groups of children. Many classic production methodologies, on the other hand, are best suited for longitudinal

studies of individual or small samples of young children. The main goal of this chapter is to provide an overview of some recent research methodologies that are likely to play an important role in furthering this area of research. Note that this chapter is not an exhaustive review of all language testing methodologies.

Infant speech perception research has focused on development in children between birth and about 20 months of age. The bulk of language production research, on the other hand, has focused on language learners at the onset of word production, that is, 18–20 months and older. (Though see work on babbling and the transition to speech, for example, Vihman et al., 2009, and other work tracking phonological development in children starting at 1 year of age, such as Demuth et al., 2006; Fikkert, 1994 and Levelt, 1994.) Infant speech perception research has demonstrated that infants know a great deal about the phonological patterns of their language long before they produce their first words (Curtin and Hufnagle, 2009; Mani, this volume). A dramatic demonstration of infants' early sensitivity to phonological patterns comes from the word segmentation literature. During the latter half of the first year of life, infants already begin using language-specific knowledge of how words typically sound to locate likely word boundaries (e.g. Johnson, 2008; Johnson and Jusczyk, 2001; Jusczyk et al., 1999; Mattys and Jusczyk, 2001). At this young age, infants appear to have highly specified perceptual representations of words. Even single segment mispronunciations in unstressed syllables are readily detected by 7.5 to 10 month olds (Johnson, 2005; Jusczyk and Aslin, 1995). Some have come so far as to argue that lexical representations are likely over-specified early in development (e.g. Houston and Jusczyk, 2000; Singh et al., 2004; though see van Heugten and Johnson, 2009).

Despite their apparent perceptual sensitivity to phonetic detail, children make systematic production errors as they produce words (Fikkert, 2007). For example, children learning Dutch make errors in producing word-initial voiced segments, such as mispronouncing *dier* 'animal' as *tier* (van der Feest, 2007). The majority of developmental speech production studies do not simultaneously examine perception, perhaps because of the assumption that child perception is adult-like. Classic studies that show that children do not accept their own mispronunciations as acceptable certainly support this view. For example, although a child says *sip* for *ship*, they will not accept *sip* for *ship* when produced by an adult (Smith, 1973). This indicates that the child can perceive the phonological differences between their (inaccurate) output and the adult target form. Another feature of child language that suggests children's perception is adult-like comes from widespread changes in their articulatory development. When a child's production abilities mature and they are able to produce segments previously produced incorrectly, children correctly produce the target segments in all contexts, without needing to rehear all instances. For example, take a child who produces the /pl/ in *please* incorrectly as *pease*. As the child

develops, they are able to correctly produce the /pl/ in *please* along with other words that contain /pl/, such as *play*. This is only possible if the child had accurately perceived the /pl/ sequence and had correctly stored representations of these targets (though see literature on lexical diffusion¹). The absence of analyses on development speech perception in early childhood has also stemmed from methodological considerations. For example, many studies of phonological development have based their analyses and conclusions on longitudinal corpus data (Ohala, 2008). While these analyses provide insights into the development of a phonological system across a single child or group of children, because they are production-based, they do not allow for analyses of children's perceptual abilities.

For these and other reasons, there is a gap in our understanding of phonological development during learners' transition to speech. Although many production studies assume that children's perception is accurate and/or that children's production errors are not the result of perceptual errors (cf. Ohala, 1999), research shows that children's perception is in fact, not entirely adult-like (Walley, 2005)². Examining both perception and production at the emergence of language production allows one to more precisely characterize phonological acquisition. For example, a study of speech production would be enhanced with a simultaneous examination of the time course of speech perception. The perception study could provide additional and crucial information about the phonetic detail of children's early phonological and lexical representations, how language is being processed, learned and represented. A study by Sundara et al. (2008) addressed these types of issues by simultaneously evaluating children's perception and production of 3rd person singular – s. Results revealed differences in performance on the two tasks: children who had a familiarity preference for grammatical sentences in the perception task had lower production accuracy scores, while children who preferred to listen to novel, ungrammatical sentences had higher production accuracy scores.

The goal of this chapter is to give an overview of a few select behavioural methodologies that are suitable to experimentally assess developmental speech perception (Anticipatory Eye Movement Paradigm) and production (Elicitation tasks and Non-Word Repetition Tasks) in children who are at the early stages of lexical development; that is, children who are approximately 20 months and older. While there is a large amount of variability when children produce their first words and many children will produce their first words before this stage, we focus on the age at which one can reasonably gather production data from young children who are being assessed in a single visit to an experimental laboratory. Children aged 20 months have an average expressive vocabulary size of around 50 words (Fenson et al., 1997, cited in Dale et al., 1998). Even though it is possible to do production studies with children younger than 20 months (Ohala, 1999), in our experience, this age is at the lower end from

which one can successfully complete experiments using elicitation and non-word repetition.

2. Methods to Assess Developmental Speech Perception

As noted above, research on developmental speech perception examines acquisition from early infancy, starting at birth. Some of the methods commonly used to research speech perception in young infants, such as the Headturn Preference Procedure (HPP) (Fernald, 1985; Kemler Nelson et al., 1995), are optimally used with younger infants. While the HPP can be used with 20 month olds, older children can become bored with this procedure, leading to high attrition rates. Moreover, this paradigm does not typically provide a very sensitive measure of individual variation, making it less than ideal for examining the relationship between perception and production. One of the most recent techniques developed to examine brain activity in neonates and young infants is near-infrared spectroscopy (NIRS) (Aslin and Mehler, 2005; Mehler et al., 2008; Meek, 2002). NIRS has been applied to examined language development (Gervain et al., 2008; Peña et al., 2003), sensory processing (Bortfeld et al., 2007), prosody (Homae et al., 2006; Homae et al., 2007) and language processing in bilinguals (Kovelman et al., 2009). This paradigm may be a promising way to test developmental speech perception in toddlers, yet it is still in development and it is yet unknown how sensitive the paradigm will be at targeting specific issues in phonological acquisition.

Testing developmental speech perception with toddlers is somewhat tricky depending on the methodology that one uses. It can be difficult to obtain explicit judgements from young children because they often do not understand the instructions and find it difficult to focus. Consider a picture-naming task used to assess discrimination of minimal pairs, such as a place of articulation contrasts between /d/ *doll* and /b/ *ball*. When asked, *Point to the doll*, 20 month olds will often simultaneously point to both pictures on the page, making it difficult to assess their perceptual abilities. Even when using a picture-pointing methodology where pictures are presented on a computer, young children may enjoy the study, but do not perform well (Parisse and Soubeyrand, 2003). In short, picture-pointing tasks and tasks that require children to provide an explicit answer may give too conservative an estimate of children's phonological knowledge (e.g. Barton, 1980; Brown and Matthews, 1997; Shvachkin, 1948/1973) as compared to on-line (implicit) measures (Fennell and Werker, 2003).

An established technique that can be used to assess speech perception in word forms with young children is the Intermodal Preferential Looking Procedure (IPLP) (Hirsh-Pasek and Golinkoff, 1996; Hollich et al., 2000; Lew-Williams and Fernald, 2007; Houston-Price et al., 2007). The IPLP has received close attention

in recent methodology overviews (Fernald et al., 2008; Gerken, 2009; Johnson and Zamuner, 2010), therefore, we will only mention this methodology in brief. In the IPLP, children are presented with trials containing two pictures shown on a visual display. For example, pictures of minimal pairs like *ball* and *doll*. Participants are asked to look at a particular picture, *Look at the ball*. Experimental analyses examine participants' proportion of looking time to the target picture (*ball*) versus the distractor (*doll*), and the relative speed to which participants look at the target and distractor. Recall that a limitation of picture pointing with very young children is that they will often point to both pictures when asked to *Point to the ball*. In the IPLP task, children may *look* at both pictures, however, the methodology also allows one to measure the proportion of looking time to either picture and the speed of looking, providing a more sensitive measure of young children's phonological knowledge.

In the field of phonological development, findings based on this methodology have had a large impact on our understanding of children's early phonological and lexical representations. In particular, results have illustrated that children's early lexical representations encoded much more phonetic detail than previous production studies have suggested (Swingley and Aslin, 2002). At 14 months, infants look longer at an object when it is produced correctly (*baby*) than misproduced (*vaby*). Though note that recent work using the same methodology has argued that early lexical and phonological representations are underspecified (van der Feest, 2007). While the IPLP is a powerful technique for investigating phonological acquisition at the onset of language production, one shortcoming is that it requires pictureable objects, from which looking times are measured. If the appropriate real-word stimuli are not familiar to young children, it may be useful to use non-word stimuli (though this methodology can also provide insights into word learning mechanisms; Byers-Heinlein and Werker, 2009). In the following section, a new paradigm called the Anticipatory Eye Movement Paradigm is described, a paradigm that does not require picturable (known) objects to test speech perception.

2.1 Anticipatory Eye Movement Paradigm

A new and promising methodology that may help researchers better understand the perception–production link is the Anticipatory Eye Movement Paradigm (AEM) (Aslin and Fiser, 2005; McMurray and Aslin, 2004). This methodology was developed to study categorization in language learning, although its foundations are in the visual expectation paradigm (VEP) (Haith et al., 1988). This methodology and variations of it have been used in different areas of cognitive development, such as spatial representations (Kaufman et al., 2006), object concepts (Johnson et al., 2003) and cognitive control abilities (Kovács and

Mehler, 2009a). It has also been used in studies of language, such as on-line sentence processing (Altmann and Kamide, 2007; Nation et al., 2003). A recent study by Kovács and Mehler (2009b) used this technique to investigate monolingual and bilingual 12-month-old infants' sensitivities to structural regularities in speech. They found that bilingual infants were better at learning multiple speech structures as compared to the monolingual infants (trisyllabic speech items with an AAB or ABA structure). The suggestion is that bilingual infants are more flexible learners and this enables them to learn multiple languages.

In AEM, infants are trained to anticipate different categories of acoustic (and/or visual) stimuli on either the left or right side of a visual display. Importantly, training stimuli unambiguously belong to 1 of 2 categories and there is typically 100% predictable relationship between what a child hears (and/or sees) and where the visual reinforcer will appear. For example, stimulus A is associated with the left side of the visual display, and stimulus B is associated with the right side of the visual display.

After infants have learned that stimuli type A leads to visual reinforcement on the left side and stimuli type B leads to visual reinforcement on the right side, they begin making anticipatory movements to the left or right side of the screen as soon as the training stimuli are presented. During test, novel stimuli are played and infants' eye movements are recorded. Test stimuli differ from the training stimuli in that they do not necessarily unambiguously belong to one category or another. For example, in one study, infants were trained on two endpoints of the *ba/pa* continuum, and then tested on new exemplars with intermediate voice onset times (VOTs) (McMurray et al., 2000). AEM has also been used to show that infants are able to identify words across changes in pitch, but not duration (McMurray and Aslin, 2004). When used in discrimination paradigms, the AEM has been shown to be more sensitive than classical behavioural tasks. For example, Albareda-Castellot et al. (in press) adapted the AEM paradigm so that all trials had visual reinforcement. Infants' anticipatory eye movements across the course of the experiment were compared to assess infants learning. Albareda-Castellot et al. found that 8-month-old Spanish-Catalan bilinguals were able to discriminate Catalan vowel contrasts, whereas studies using familiarization methodologies with the same aged and language background infants did not find discrimination (Bosch and Sebastián-Gallés, 2003).

AEM studies can be designed in a number of ways. As stated before, studies typically have a training phase and a test or generalization phase. Experiment designs may require a predetermined number of training trials before the test phase (McMurray and Aslin, 2004). Alternatively, the length of the training phases may be tailored to individual participants. This is due to the recent development of software that analyses infants' anticipatory looking over the

course of training. For example, programs can calculate an on-line measure of the number of correct trials with correct anticipatory looking. Based on these measures, it is possible to stipulate that once participants have reached a preset criterion of looking, they proceed to the test phase where generalization trials are presented with novel and potentially ambiguous stimuli. As with the IPLP, the dependent measure can vary from the proportion of looking time to the either side of the screen, the relative speed to which participants look to either side of the screen or with time course analyses. Analyses of the training phase and test phase will depend on the experiment design. Importantly, analyses of the training phase should examine whether participants make the correct anticipatory eye movements. Analyses of test trials will show how individual participants or groups of participants categorize or generalize new stimuli in comparison to the training stimuli.

The AEM is adaptable and is not limited to any specific age or linguistic groups; however, it is not appropriate for infants under 2 months, as controlled eye movement is still developing around this time. In our work, we have begun using this paradigm with 12 to 20 month olds, and are finding the task to be well suited to this age range. AEM studies are best presented using eye-tracking systems because they allow for precise measurements and timing of infants' eye movements. Moreover, with the use of eye-tracking techniques, the procedure easily allows for reaction time measures. A major disadvantage in using AEM is the relative newness of the paradigm. Standard experimental designs and analyses have not yet been established. Although few studies have been published using this paradigm with phonological development, AEM studies have a long history in other areas of cognitive development.

3. Methods to Assess Developmental Speech Production

Experimental research in phonological speech production has received less attention than in developmental speech perception. Typically, production studies have focused on evaluating the development of individual or groups of children's phonological systems. This approach stems in part from Smith's landmark study (Smith, 1973), in which he provided a comprehensive analysis of his son's phonological development. Large-scale corpora studies are labour intensive, though database sharing through CHILDES (MacWhinney, 2000) and the development of corpora tools such as PHON, an open-source program to manage phonological corpora (Rose et al., 2006), are making it attainable for researchers to investigate phonological development using large-scale, longitudinal corpora. For a review of issues pertaining to the development of corpora suitable for examining language development see Demuth (2008).

There are a number of techniques that can be used to assess phonological development in production studies with children. For example, word-games require participants to manipulate parts of words allowing researchers to examine participants' knowledge of words' segmental and syllable structure (Fallows, 1981; Treiman and Danis, 1988; Zamuner and Ohala, 1999). To illustrate this, take a pause-insertion task. Here participants are asked to repeat words with a pause between syllables, such as *tiger* is repeated as *tig(pause)er*, *ti(pause)ger* or *tig(pause)ger*. Analyses of children's responses determine whether participants associate the medial 'g' with the first syllable, second syllable, or as ambisyllabic. The association (syllabification) of the medial consonant is influenced by factors such as vowel quality and stress. Another task used with young children is the Wug Task, which was designed to assess learners' productive morpho(phonological) knowledge (Berko, 1958). Participants are given novel words (*This is a wug.*) and asked to produce a morphological variant (*Now there is another one. There are two of them. There are two ____.* Answer: *wugs.*). While methodologies using word-games and other types of manipulations are informative, they are also difficult for toddlers to complete. Therefore, we focus on two methodologies that have been successfully used with children under the age of two: elicitation and non-word repetition. We also point the reader to Blom and Unsworth (2010), for a discussion of issues in running experimental language studies with children.

3.1 Elicitation Studies

When using corpus data, a common drawback is that in a typical recording session of spontaneous, naturalistic data, children will not produce enough words (or any words) containing the phonological structure under investigation. Take an example of naturalistic data on young children's early morpho-phonological voicing alternations in Dutch from the CHILDES database (Fikkert, 1994; Levelt, 1994). In Dutch, voicing alternates between [t-d] in pairs like [brɒt-brɒdɛn] 'bread-breads'. Although the database contains over 20,000 utterances collected from 12 children acquiring Dutch, there are only 9 types that have the targeted voicing alternations (Kerckhoff, 2007: 132). To address this limitation in using naturalistic data, one of the simplest tasks to use with young children is elicitation (Ohala, 2008). In this methodology, children are prompted for words that are targeted for their phonological characteristics. Elicitation tasks are based on the assumption that the patterns of correct and incorrect productions reflect learners' phonological knowledge, language acquisition mechanisms, and language representations. Using elicitation task, researchers have examined a wide range of phonological patterns. For example, studies

have investigated learner's knowledge of segmental structure, using words with different clusters to examine children's error substitutions (Kirk, 2008), and studies have examined the acquisition of prosodic structure, using multisyllabic words to study truncation patterns (Kehoe and Stoel-Gammon, 1997).

There are a variety of tasks that can be designed to elicit productions of specific words. A very successful method is the picture-naming task. In our experience, this is more successful when pictures are presented on the computer as children find it very engaging. It is also possible to intersperse entertaining video trials, to maintain their attention, which will vary depending on the age of a child. Another major advantage of using a computer is that it keeps children in a specific location within the room, allowing for more reliable audio and video recordings. This can be essential for later acoustic analyses and/or transcriptions. Children will regularly place their hands in their mouth, and if these are captured on video, one is able to exclude these productions. See the CHILDES website for a complete discussion on how to make audio and video recordings (www.childes.psy.cmu.edu).

Many issues can influence children's production, and it is essential to consider what factors to control in the set of experimental stimuli and methodology. For example, recent work by Edwards and Beckman (2008) compared children's production of word-initial segments in Cantonese, English, Greek, and Japanese. Children's segmental accuracy was effected by the frequency of the neighbouring phoneme context (initial segments were more accurate when followed by a frequent vowel than an infrequent vowel), word length (initial segments were more accurate in shorter words than longer words), and prosodic factors such as word stress and pitch accent (initial segments in Japanese are produced more accurately when they are in syllables with high tone than low tone). Other factors such as utterance position can influence the accuracy of children's production. For example, the 3rd person singular morphemes (-s in *looks*) is produced more often by children in utterance-final than utterance-medial position (Song et al., 2009). Another factor is word frequency; children are better at producing high frequency words and frequently occurring phonological patterns (Zamuner, 2003). Therefore, before deciding on the set of experimental items, it is often useful to calculate the distribution of phonological patterns in the ambient language in an appropriate corpus (e.g. a corpus of child directed speech). Controlled experimental studies can help factor out some of these influences on children's productions, which are difficult to control for in analyses of corpora data.

Elicitation tasks can be very successful with young children. The only limitation is that participating children must be comfortable enough in the experimental setting to produce speech. Young children are often shy in new situations; therefore, it is useful to first use a task or play a short game that does

not require verbal responses. In some cases participants will not produce any spontaneous responses, thus, it is useful to prepare an imitation task for the same stimuli. This may be part of a between-subjects condition where imitation can be compared to spontaneous productions. Imitation can be informative because it can help interpret elicited productions. For example, using a picture-naming task, Kerkhoff (2007) found that 3½ year old Dutch-learning children do not reliably produce a medial voicing contrast in certain morphological conditions. In bi-morphemic contexts, children produce both medial 't' in *petten* 'caps' and medial 'd' in *bedden* 'beds' as 't', *petten* and *betten*, respectively. In mono-morphemic conditions, children accurately produce a voicing contrast, the medial 't' in *water* and the medial 'd' in *poeder* are produced correctly. However, when the same children were tested on an imitation task, a different result was found; children were equally accurate at imitating medial 't' and medial 'd' in both morphological contexts. This provides evidence that children at this age do not have difficulty producing a medial voicing contrast, but that other lexical factors influenced their performance in the picture-naming task.

When designing an elicitation task, there will be a limit on the number of words that you can reasonably expect children to produce in a single experimental session, which will change depending on the specific age or age range of the participants. To cope with these limitations, one could use a between-subjects design or have multiple testing sessions with the same child. When the study includes very young children, extensive piloting is helpful to determine whether children are likely to be familiar with the items and pictures. To help solicit items, specific frames should be prepared and used with all children for consistency. It may also be necessary to include filler items. Research has shown that over sampling of a specific word shape or phonological pattern may prime children to produce words in a specific way, lead to a higher number of errors than found in spontaneous speech. For example, an analysis of children's overgeneralization errors found that children produced more overgeneralization errors in an experimental setting than found in spontaneous speech. In the experimental study, children were provided with a verb stem and asked to produce it in the past tense, as in 'I will drink my milk. I already ____ my milk.' (Kuczaj, 1978; Marcus et al., 1992). In this context, children were more likely to overgeneralize 'drinked' than in spontaneous speech samples. The use of fillers can help offset these potential priming effects.

Once the data are collected, they will need to be transcribed and coded. Protocols for phonetically transcribing child production data vary across researchers and experimental labs. For example, variation is found in the use of trained phonetic transcribers who may or may not be native speakers of the language, how naïve the transcribers are to the experiment purpose, and the amount of transcribed data that is checked for inter-transcriber reliability. The

level of phonetic transcription also varies (narrow versus broad) depending on the goals of the study, though it should be noted that there is less reliability across transcribers on narrow phonetic transcriptions. Many of these issues have important implications for transcribed data. See Edwards and Beckman (2008) for a recent discussion on how to deal with issues of phonetic transcriptions, such as influences of transcribers' native-language bias in their transcription of non-native language phonemes. Some of these issues can be addressed by including acoustic analyses of children's production. These types of analyses may reveal covert contrasts that exist in children's speech, but that are not audible to adult transcribers (Buder, 1996; Scobbie et al., 2000). Alternatively, acoustic analyses may be the primary way in which the data are coded, depending on the study's goals.

Data analyses will depend on the goals of the study. An experimental goal may attempt to determine the sound patterns that children are able to produce at different stages. For example, if one is investigating whether there are prosodic interactions in children's segmental productions, one may compare children's accurate production of word-initial versus word-final segment, or the different types of segments produced accurately in these positions (such as targets with labial, coronal or dorsal place of articulation) (e.g. Beers, 1995). It may also be informative to do an error analysis, to see whether children's errors show specific patterns of results. For example, when young children first attempt consonant clusters, they typically will delete one segment. Analyses may examine whether children reduce the first or second member of the cluster, for example, does *pretty* reduce to *pitty* or *ritty*. These types of analyses have revealed that children's cluster reductions adhere to the well-formedness of syllable sonority (Jongstra, 2003). Alternatively, analyses may examine participants' segmental substitutions. For example, these types of analyses have revealed that children's segmental substitutions result in clusters that share place or manner of articulation, for example, *ducks* becomes *duts* (Kirk, 2008).³ Other types of error analyses may examine whether children's substitution errors are towards the more frequent segments or more frequent phonological features of the language. Also it may be possible to collect reaction time measures from children's productions using a voice key, which provides an automatic and electronic measure (rather than an off-line, manual evaluation) of the time between stimulus presentation and the onset of speech production (Tyler et al., 2005). In this case, one might compare reactions for children's correct productions versus incorrect productions of initial segments.

An unavoidable problem that is frequently encountered when working with young children is empty data cells. That is, young children may not produce a specific item for various reasons. This can usually be addressed by taking the proportion of correct responses for a condition. For example, consider an experiment designed to examine children's production abilities of place of

articulation. In this hypothetical study, there are four words in the stimulus set to evaluate how children produce word-initial /b/: *bear*, *book*, *boat* and *bed*. Imagine a participant who correctly produces /b/ in *bear* and *book*, misproduces the /b/ in *boat*, and does not give a response for *bed*. For this participant, their accuracy score on word-initial /b/ would be 0.66 (2 out of 3 words with word-initial /b/ were produced correctly).

Although there are many advantages to using real-word stimuli, there are also potential limitations. With real-word stimuli, it is often difficult to find enough appropriate items that are picturable or familiar to young children. The targeted sound pattern may occur most often in verb or adjectives, making it difficult to elicit spontaneous productions of these words as compared to nouns. For example, there are very few common English words known to children that end in /v/ as in *love*, whereas it is very easy to find nouns that end in /t/, as in *cat*. The frequency of individual lexical items also has an effect on how accurately children produce words. For example, Zamuner (2003) found that young children are quite accurate at producing final 'th' in the word *bath*, even though 'th' is a late acquired segment and is difficult for young children to produce. Other processes such as lexical diffusion may influence children's performance. Lexical diffusion refers to how sound changes evolve in children's developing phonological stems. Children's acquisition of fricatives has shown different patterns of development depending on word frequency and where the target fricative occurred within a word's position (Gierut and Storkel, 2002). Given these types of potential limitations of using real-word stimuli, it can be hard to design controlled experiments to test phonological development of specific phenomena.

3.2 Non-Word Repetition Tasks

One methodology that circumvents these problems is the non-word repetition task (NWRT). The NWRT has been widely used to assess the development of children's phonological and lexical representations, speech perception, articulation and memory (see Coady and Evans, 2008 for a recent review). For example, parallel to the studies that have used elicitation to study the acquisition of consonant clusters, Ohala (1999) found in a NWRT that children's reductions of initial and final clusters were predicted by sonority. Initial clusters were reduced to produce a rise in sonority ([stɪg] → [tɪg]), whereas final cluster reduction more often led to a minimal sonority descent ([dʌst] → [dʌs]). Similarly, studies looking at prosodic acquisition have examined truncation patterns in young children's production of multisyllabic words (Gerken, 1994).

In NWRTs, children are simply asked to repeat non-words that are controlled for various phonological properties (or other types of properties). Typically the

methodology is used with children over the age of 3, although studies have been successful with participants under 2 years of age (Zamuner, 2003). NWRTs are ideal for children at the beginnings of language production because they capitalize on ‘echoism’ (Jespersen, 1922/1964) or ‘echolalia’ (Guillaume, 1926/1971). That is, children imitate speech. In a research setting, young children who are presented with non-words will spontaneously repeat them, with needing explicit instruction to do so. As with the picture-naming task, we have had the most success presenting pre-recorded non-words over a computer. This also controls the acoustic properties of the non-word stimuli, so that all subjects are presented with the same tokens.

Like studies using elicitation, NWRTs require careful attention to stimuli design. The frequency of the sound components of non-words is an important factor to consider, as children are more accurate at producing frequent segments and segmental combinations (Coady and Evans, 2008; Munson et al., 2005). Moreover, studies have revealed that young children are more accurate at producing the same sound depending on the frequency of the non-word components. Zamuner (2009) found that children are better at producing word-initial /p/ in non-words composed of high frequency patterns than low frequency patterns. A standard way in the field is to control non-word stimuli for their phonotactic probabilities, that is, the likelihood that a sound has to occur in a given word environment (Storkel, 2004). Stimuli are also typically controlled for their neighbourhood densities, which is a measure of the number of similar sounding words in the lexicon.

Many of the same considerations described in the previous section on elicitation apply to data transcription, data coding and data analyses for NWRt. A unique consideration in studies using non-word production is the treatment of real-word responses. Young children may often produce a real word in response to a non-word. For example, in Zamuner, Gerken and Hammond (2004), a typical error in children’s production of the non-word *bome*, was to substitute the final ‘m’ as ‘n’, producing *bone*. It is possible that children misperceived *bome* as *bone*, though it is also possible that children perceived it correctly and made a segmental production error. To address this problem, we have typically excluded real-word answers when they are words known to young children – that is, if the real word occurs in a corpus of speech known to young children. In this case, children’s productions such as *bone* would be excluded, whereas a production of *bode* would not be.

4. Conclusion

In this chapter we have described methodologies suitable for investigating the simultaneous development of speech perception and production. The

perceptual task we have focused on is the Anticipatory Eye Movement Paradigm. The two production tasks we have focused on are word elicitation (picture-naming task) and the NWRT. There are two broad approaches to combining these sorts of methodologies to examine the development of perception and production in tandem. To illustrate, take the case of VOT acquisition. One approach would be to examine VOT perception and production in the same child. For example, following McMurray et al. (2000), one could devise a perception study using the AEM to examine a child's language-specific voicing contrasts. In other words, to design a study that would establish the VOT boundary for an individual child. This same child could then be tested on a production study using real-words or NWRT. Comparisons of the same child's system could be made to determine whether individual children's perceptual and production systems align. Another approach would be a similar type of study, but to compare group data on perception and production. Examining developmental speech perception and production in concert is an upcoming challenge for researchers. Yet the future is bright for research exploring the relationship between speech perception and speech production. We hope that the methodologies described here will provide useful tools for researchers interested in these overlapping and intertwined areas of language development.

5. Notes

1. Studies show that sound-changes in children's phonological system depend on many factors, such as the type of sound change, word position and word frequency (Gierut, 2001; Gierut and Storkel, 2002).
2. Note that by adult-like, we do not mean 'perfect perception'. Adult's perception is not perfect or free of errors. Adults are known to make the occasional 'slip of the ear' (e.g. Bond, 1999) as well as the occasional 'slip of the tongue' (Cutler, 1982).
3. Two primary methods used to evaluate children's productions are the independent analysis and relational analysis. Independent analyses measure children's productions independent of the adult target without consideration of whether children produce the adult target correctly. Relational analyses measure children's productions as they relate to the adult target form (Stoel-Gammon, 1985; Stoel-Gammon and Sosa, 2006).

This page intentionally left blank

Part II

RESEARCH ISSUES

This page intentionally left blank

3 Features

*Bert Botma, Nancy C. Kula and
Kuniya Nasukawa*

Chapter Overview

1. Introduction	33
2. Feature Theory	36
3. Element Theory	42
4. Voicing and Nasality in Zoque	50
5. Zoque Analysis	55
6. Conclusion	62
7. Acknowledgements	62
8. Notes	62

1. Introduction

A fundamental assumption in the study of language sounds is that the segmented parts of the speech continuum (speech sounds) can be decomposed into smaller properties. Support for these smaller properties comes first and foremost from the notion of ‘natural class’. Natural classes are composed of sounds which participate in the same phonological process(es). Sounds which form a natural class are assumed to share some properties or combination of properties to the exclusion of other sounds. In English, for example, when the first consonant of a C_1C_2 sequence is nasal, it is typically homorganic (shares the place of articulation) with a following stop, as in *temper* [tɛmpə], *banter* [bæntə], *anger* [æŋgə]. On the other hand, if the left-hand consonant of a C_1C_2 sequence is not nasal, then it is not necessarily homorganic with the right-hand consonant, as in *gulp*, *casket*, *act*. We can say, therefore, that sounds which are homorganic with a following consonant form a class as compared to sounds which are not: the former share the property of being nasal, the latter of being oral. More generally, we can say that the number of natural classes to which a particular sound is affiliated is a diagnostic for the number of distinctive

properties of which the sound is comprised (e.g. Chomsky and Halle, 1968; Hyman, 1975; Lass, 1984; Hall, 2007).

Unlike phonemes, which are constructed on a language-specific basis, the properties inherent in speech sounds are generally thought to function as universal 'primes' which form part of the phonological component of the language faculty. These properties are referred to by such labels as 'features' (Jakobson et al., 1952; Chomsky and Halle, 1968), 'elements' (Kaye et al., 1985; Harris and Lindsey, 1995), 'components' (Anderson and Ewen, 1987; van der Hulst, 1995) and 'particles' (Schane, 1984). With respect to nasality, for example, a number of different kinds of primes have been proposed. In traditional feature theory the relevant prime is the bivalent (equipollent, binary-valued) feature [\pm nasal]. In element-based approaches, on the other hand, which assume that primes are monovalent (privative, single-valued), nasality has been variably represented as |N| (e.g. Kaye, 1989; Harris, 1990; Harris and Lindsey, 1995; Nasukawa, 2005; Nasukawa and Backley, 2008; Backley and Nasukawa, 2009a and b; see also Davenport, 1994) or |L| (e.g. Kula and Marten, 1998; Kula, 1999, 2002; Ploch, 1999; Botma, 2004; Botma and Smith, 2006, 2007).

A theory of phonological primes must be capable of representing both the distinctive properties of segments and the various operations and processes affecting them. Not unsurprisingly, it has proved a challenge to adequately account for these aspects. A general problem faced by all theories is overgeneration in terms of the number of primes and the number of possible operations. In this respect an important development has been the position that at least some primes are monovalent (privative, single-valued). A case in point is nasality. The observation that nasals in a language like English share their place of articulation with a following stop is an argument for taking either [\pm nasal] or [+nasal] to be phonologically relevant. However, since there appear to be no compelling examples of natural class behaviour of non-nasals, that is, of all oral sounds, the evidence tips the balance in favour of a single-valued feature [\pm nasal]. The point is that a binary-valued feature [\pm nasal] overgenerates, with [-nasal] predicting a segment class that is phonologically irrelevant (see, for example, Rice and Avery, 1989 and Steriade, 1993a for arguments to this effect).

The position that all phonological primes are privative has been defended most extensively by proponents of Dependency Phonology (DP) (Anderson and Jones, 1974; Anderson and Ewen, 1987), Government Phonology (GP) (Kaye et al., 1985), Element Theory (ET) (Harris and Lindsey, 1995), Particle Phonology (Schane, 1984) and Radical CV Phonology (van der Hulst, 1995). While this position is in principle more restrictive, the question is whether it can be maintained in the face of empirical data. The phonological literature suggests that some primes, for example, those referring to place of articulation properties, are more amenable to a privative analysis than others, such as voicing, continuancy and tongue-root advancement/retraction. In addition, Rice (2009)

observes that privativity often leads to added complexity in some other domain of the grammar. An example of this might be the representation of nasality in the ‘standard’ DP model of Anderson and Ewen (1987), where nasal manner is represented by $\{V;C\}$ (i.e. ‘vocalicness governs consonantality’) for nasal consonants and by the component $|n|$ for nasalized sounds. A similar representation is used in some versions of Feature Geometry, where nasality is represented by both the Soft Palate node and the feature [nasal] (see, for example, Sagey, 1986; Halle, 1992; Piggott, 1992). It is not immediately clear whether such accounts are more restrictive than one in terms of a single, binary-valued feature $[\pm\text{nasal}]$.

Despite these concerns we will assume the basic appropriateness of a privative approach to primes in this chapter, at least as a working hypothesis. We are supported in this by recent work on assimilation and harmony, which shows a renewed interest in the issue of feature privativity (see McCarthy, 2009).

Aside from privativity, another attempt at theoretical parsimony has been the reduction of the number of primes themselves. The frameworks of Dependency and Government Phonology offer articulated attempts at integrating these two strategies into models of segmental structure. These approaches assume not only that primes are privative, but also that one and the same prime can have a different (but related) phonetic exponence depending on its position in the phonological structure. This reduces the overall number of primes required to express lexical contrasts and thus simplifies the phonological grammar.

Our goal in this chapter is to defend one particular attempt at theoretical parsimony in the segmental domain. Combining insights from DP, GP and ET, we show that there are good arguments for assuming privative features and in particular a single element, viz. $|L|$, which represents both voicing and nasality. In the first part of the chapter we discuss different approaches to feature representations. In the second part we defend a privative, reductive approach in terms of $|L|$ by investigating a case study of voicing and nasalization in Zoque, a Mixe-Zoque language of Mexico. Although the Zoque facts do not seem immediately amenable to such an approach, we show that they follow naturally from the assumption that voicing and nasalization operate at different levels in prosodic structure.

The chapter is organized as follows. In Section 2 we give a short overview of feature theory and discuss some problematic aspects of this framework. In Section 3 we offer a general introduction to the notion of dependency as used in DP and some versions of ET, and outline our assumptions regarding the representation of voicing and nasality. This sets the stage for our discussion of Zoque. We present the relevant Zoque data in Section 4 and offer our analysis of them in Section 5. Section 6 concludes.

2. Feature Theory

2.1 An Overview of Distinctive Features

Distinctive features were first recognized as taxonomic properties for categorizing phonemes in earlier structuralist theories of phonology (Jakobson, 1939a [1962], Trubetzkoy, 1939; see also Scheer, this volume). Jakobson et al. (1952) and Jakobson and Halle (1956) developed the idea that phonemes are composed of bundles of features, with the latter forming the minimal units of lexical contrast. Jakobson and Halle employed 12 features, whose phonetic exponence was defined in primarily acoustic terms and only secondarily in articulatory terms. For example, nasality, formalized by the feature value [+nasal], was defined in acoustic terms as involving the presence of additional formants, coupled with a reduced intensity of existing formants. The articulatory correlate of this is a lowered velum, with the presence of (voiced) nasal airflow. The complete list of Jakobson and Halle's features is given in Table 3.1, together with their acoustic and articulatory correlates.

This feature system accounts for the observation that different articulatory gestures may produce similar acoustic effects. For example, lip rounding, pharyngealization and retroflexion are partially subsumed under the feature [flat], while lip rounding and velarization are partially subsumed by [grave]. The system also employs the same features for the description of both consonants and vowels. For example, [grave] characterizes not only labial and velar consonants but also back vowels.

All features in Jakobson and Halle's framework involve polar oppositions (with feature judged redundant or irrelevant left blank). For example, using the features in Table 3.1, English *tent* can be represented as a sequence of four sets of features, as in Table 3.2, where we disregard the irrelevant features [checked], [flat] and [sharp].

In this approach the homorganicity of *n* and *t* can be accounted for in terms of [-compact] and [-grave], which are shared by the two segments. Similarly, in *anger* the nasal and the following stop share [+compact] and [+grave], while in *temper* they share [-compact] and [+grave]. [±compact] and [±grave] thus form a class with respect to the homorganicity of nasal-stop clusters. Notice though that the feature system in Table 3.2 is unable to capture this feature grouping; the combination of [±compact, ±grave] is an essentially random combination of features.

The Sound Pattern of English (SPE) (Chomsky and Halle, 1968), the most influential post-Jakobsonian work on features, employs a different set of features to describe the homorganicity of nasal-stop clusters, including [±coronal], [±anterior], [±high] and [±back]. In SPE, features are defined in articulatory terms; for example, [+nasal] is defined simply as involving a lowered velum.¹

Table 3.1 Jakobson and Halle's distinctive features

<i>Feature</i>	<i>Opposed to</i>	<i>Acoustic description</i>	<i>Articulatory description</i>
<i>Sonority features</i>			
1 Vocalic	Nonvocalic	Sharply defined formant structure	Voiced, with free passage of air through vocal tract
2 Consonantal	Nonconsonantal	Low tonal energy	Obstruction in vocal tract
3 Nasal	Oral	Additional formants and less intensity in existing formants	Coupling of nasal cavity
4 Compact	Diffuse	Energy concentrated in central area of spectrum	High ratio of front resonance chamber to back
5 Abrupt	Continuant	Interruption or abrupt transition	Rapid closure and opening of vocal tract
6 Strident	Non-strident (Mellow)	High intensity noise	'Rough-edged' effect at point of articulation
7 Checked	Unchecked	High rate of energy discharge	Glottalized
8 Voiced	Voiceless	Periodic low frequency excitation	Vocal fold vibration
<i>Protensity features</i>			
9 Tense	Lax	High energy with greater spread across spectrum and longer duration	Greater deformation of vocal tract from its rest position
<i>Tonality features</i>			
10 Grave	Acute	Energy concentrated in lower frequencies	Peripheral (towards front or back of vocal tract: e.g. back vowels, labial and velar consonants)
11 Flat	Non-flat (Plain)	Downward shift or weakening of upper frequencies	Narrowed aperture (e.g. lip rounding, pharyngealization, retroflexion)
12 Sharp	Non-sharp (Plain)	Upward shift of upper frequencies	Reduced oral cavity and widened pharynx

Source: Clark et al., 2007: 442; based on Jakobson and Halle 1956.

The complete list of SPE features, with their articulatory correlates, is given in Table 3.3.

Subsequent work in feature theory has generally retained the SPE approach to feature definitions, though not all the features themselves have been retained (for a recent overview, see Hall, 2007). One reason for this might be that

Table 3.2

	<i>t</i>	<i>e</i>	<i>n</i>	<i>t</i>
Vocalic	-	+	-	-
Consonantal	+	-	+	+
Nasal	-	-	+	-
Compact	-	+	-	-
Abrupt	+	-	-	+
Strident	-	-	-	-
Voiced	-	+	+	-
Tense	+		-	+
Grave	-	-	-	-

articulatory movements can be relatively easily observed from the altering positions of the speech organs, whereas the acoustic properties of speech sounds are much harder to observe – although in recent years acoustic data have become more easily obtainable with the advent of software such as *Praat* (Boersma and Weenink, 2009). It has also been claimed that articulatory feature descriptions are appropriate since articulation is the ultimate substance of speech (e.g. Clark et al., 2007: 374), a point to which we return shortly.

With a subset of the features in Table 3.3, English *tent* can be represented as shown in Table 3.4.

Thus *n* and *t* share the feature values [+coronal, +anterior, –high, –back] (the nasal and the following stop in *anger* are [–coronal, –anterior, +high, +back]; in *temper* they are [–coronal, +anterior, –high, –back]). In SPE, these features are all listed under the label ‘cavity features’; though notice that in this system, too, there is no explicit grouping of them as place of articulation features. What is required in addition to features, therefore, is a way of identifying what makes them function as a group as well as a way of excluding unattested feature groupings. This is the main motivation behind the framework of Feature Geometry, which we consider next.

2.2 Feature Geometry: Hierarchical Relations of Features

The framework of Feature Geometry (FG) is motivated by place feature groupings of the kind in Table 3.2 and Table 3.4, as well as by various other phonological processes such as dissimilation, lenition and fortition, which also show that certain features typically function as a class. FG formalizes feature groupings in terms of hierarchical structure. A number of different proposals can be found in the literature (see, for example, Clements, 1985; Sagey, 1986; McCarthy,

Table 3.3 Chomsky and Halle's distinctive features

<i>Feature</i>	<i>Articulatory description</i>
<i>Major class features</i>	
1 Sonorant	Produced with vocal tract cavity configuration in which spontaneous voicing is possible
2 Vocalic	Constriction does not exceed that of high vowels, and position of vocal folds allows spontaneous voicing
3 Consonantal	Radical obstruction in mid-sagittal region of vocal tract
<i>Cavity features</i>	
4 Coronal	Produced with blade of tongue raised from neutral position
5 Anterior	Produced with obstruction in front of palato-alveolar region
6 High	Tongue body above neutral position
7 Low	Tongue body below neutral position
8 Back	Tongue body retracted from neutral position
9 Round(ed)	Narrowing of lip orifice
10 Distributed	Constriction extends for some distance along direction of airflow
11 Covered	Pharynx walls narrowed and tensed and larynx raised (in vowel production)
12 Glottal constriction	Constriction of vocal folds
13 Nasal	Lowered velum
14 Lateral	Lowered side(s) of mid-section of tongue
<i>Manner of articulation features</i>	
15 Continuant	Primary constriction in vocal tract does not block air flow
16 Instantaneous release	Instantaneous release (of stops)
17 Velar(ic) suction	Velar closure producing suction (clicks)
18 Implosion	Glottal closure producing suction (implosives)
19 Velar(ic) pressure	(Velar closure producing pressure — no evidence of use in language)
20 Ejection	Glottal closure producing pressure (ejectives)
21 Tense	Deliberate, accurate, maximally distinct articulation (of supraglottal musculature)
<i>Source features</i>	
22 Heightened subglottal pressure	Tenseness in subglottal musculature producing greater subglottal pressure
23 Voiced	Vocal fold vibration (induced by appropriate glottal opening and airflow)
24 Strident	Turbulence (in fricatives and affricates) caused by nature of surface, rate of airflow and angle of incidence at point of articulation
<i>Prosodic features</i>	
25 Stress	
26 Pitch (high, low, elevated, rising, falling, concave)	
27 Length	

Table 3.4

	<i>t</i>	<i>e</i>	<i>n</i>	<i>t</i>
Sonorant	-	+	+	-
Vocalic	-	+	-	-
Consonantal	+	-	+	+
Coronal	+	-	+	+
Anterior	+	-	+	+
High		-	-	-
Low		-	-	-
Back	-	-	-	-
Round	-	-	-	-
Nasal	-	-	+	-
Continuant	-	+	-	-
Tense	+	+	-	+
Voiced	-	+	+	-

1988; Halle, 1992, 1995; Clements and Hume, 1995; Halle et al., 2000); the geometry in Figure 3.1 is based on the influential version of McCarthy (1988).

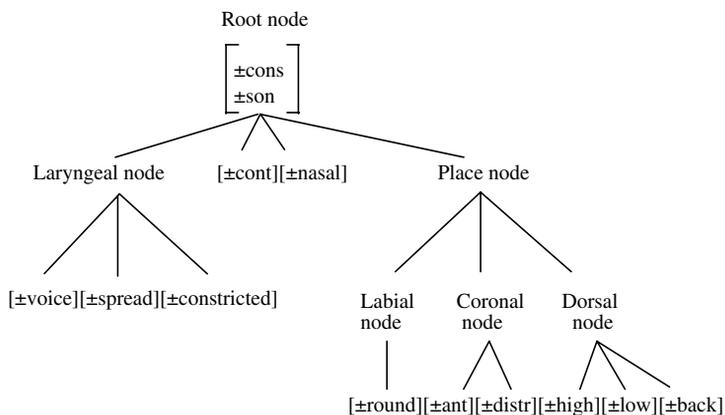


Figure 3.1 A feature tree

The highest node in the featural organization is the Root node, which consists of the major class features [±consonantal] and [±sonorant]. The Root node dominates the two ‘class’ nodes Laryngeal and Place, as well as the features [±continuant] and [±nasal]. The Laryngeal node dominates features such as [±voice], [±spread glottis] and [±constricted glottis]. The Place node

dominates three distinct articulatory nodes: Labial (subsuming [round]), Coronal (subsuming [anterior] and [distributed]) and Dorsal (subsuming [high], [low] and [back]). Most of these assumptions are fairly uncontroversial within FG; see, for example, McCarthy (1988) and Clements and Hume (1995) for discussion.

Given Figure 3.1, the homorganicity of nasal-stop clusters can be interpreted as involving the sharing of the Place node together with the features it dominates, as shown in Figure 3.2.

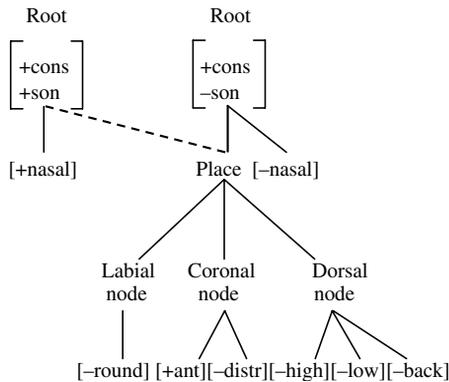


Figure 3.2 Nasal-stop place assimilation (as in *temper*)

In ‘dynamic’ processes of nasal place assimilation, that is, those involving morphological boundaries, as in *te[n] + [b]irds → te[m b]irds*, *te[n] + [k]ids → te[n k]ids*, assimilation involves the leftward spreading of the Place node from the word-initial stop to the final nasal of the preceding word. Again, note that since the Place node dominates the various Articulator nodes, spreading of Place entails spreading of all the Articulator nodes together with the features dominated by them. We can say, therefore, that in FG there is a relation of ‘dependency’ between dominating and dominated nodes.

Most of the dependency relations assumed in FG follow from the articulatory definition of the features involved. As Ewen (1995: 581) notes, ‘the interpretation of dependency relations represents an attempt to formalize the constraints on human articulators’. However, such dependency relations are more straightforward for some features than they are for others. For example, while a feature like [distributed] is uncontroversially dominated by Coronal, the location of [nasal] is much less clear. It has been proposed that [nasal] is located under the Manner node (Clements, 1985), the Peripheral (or Laryngeal) node (Hayes, 1986), the Root node (McCarthy, 1988), the Supralaryngeal node (Trigo, 1993), the Soft Palate node (Sagey, 1986), the Spontaneous Voice node (Rice and Avery,

1989), under either the Soft Palate or Spontaneous Voice node (Piggott, 1992), or under both (Tourville, 1991). Each of these proposals has its merits but, as Humbert (1995: 13) notes, they cannot all be correct.

The variety of proposals regarding the location of [nasal] suggests to us that the structural principles underlying FG are not sufficiently constrained. In this framework the motivation for a particular featural organization comes from articulatory phonetic considerations and from the behaviour of features in phonological processes. However, if such evidence is unavailable or inconclusive, there are no independent principles that help constrain the range of possible structures.

3. Element Theory

3.1 Differences between Elements and Features

We now turn our attention to the set of universal primes called ‘elements’, which have been proposed as an alternative to features. The idea of element-like components has its roots in the |a, i, u| approach to vowel place as introduced in the DP framework of Anderson and Jones (1972). The first work that explicitly offers a theory of elements, that is, ET, is Harris and Lindsey (1995), with some of the proposals there anticipated in such works as Kaye et al. (1985); Rennison (1986); Harris (1990); Charette (1991); Backley (1993); Nasukawa (1995) and Brockhaus (1995). Subsequent work in ET includes Cyran (1997); Backley and Takahashi (1998); Cabrera-Abreu (2000); Ingleby and Brockhaus (2002); Kula (2002); Botma (2004); Scheer (2004); Nasukawa and Backley (2005, 2008) and Backley and Nasukawa (2009a and b).

Elements differ from orthodox features in a number of respects, most obviously in the number of primes involved. Feature theories typically employ more than 20, sometimes more than 30 features. ET assumes a significantly smaller number of elements. For example, ‘standard’ GP (e.g. Charette and Göksel, 1994; Kula and Marten, 1998; Kula, 1999) as well as the ET versions of Botma (2004), Nasukawa and Backley (2005, 2008) and Backley and Nasukawa (2009a and b) recognize no more than six elements (we consider these in Section 3.3). From the general perspective of theoretical parsimony, a small set of primes is preferable since it minimizes the risk of overgeneration.

An important difference between elements and SPE-type features concerns their phonetic exponence. According to the view espoused in SPE, phonology is concerned primarily with speech production rather than perception (for a similar view, see Liberman and Mattingly, 1985; Fowler, 1986; Browman and Goldstein, 1989). ET rejects this in favour of an approach where phonological objects are associated first and foremost with properties of the acoustic signal,

in line with the work of Jakobson (see Section 2.2). Support for a hearer-oriented perspective comes among other things from language acquisition. As Nasukawa and Backley (2008: 4) note, studies of early language acquisition show that speech perception is prior to and independent of speech production (see also Mani, this volume). It is generally assumed that infants begin to acquire language by first perceiving adult input forms; then, on the basis of this input, they build mental representations, which serve as the beginning of their native lexicon. Only later do children begin to reproduce these stored forms as spoken language. This implies that speech perception is an indispensable stage on the acquisition path, and a prerequisite for successful acquisition. A further argument for a perception-based approach to primes comes from phonological interactions between sounds that are acoustically but not articulatorily related, such as rounded segments and labials (and sometimes velars), a point that is also made in Smith (1988). Consider, for example, the change *k^w > p in Proto-Indo-European (e.g. Campbell, 1974).

Nasukawa and Backley (2008: 37) assert that when decoding speech listeners naturally seek out linguistic information: listeners ignore most of the incoming acoustic stream, focusing only on the specifically contrastive properties contained in it. This leads Nasukawa and Backley to argue that the mental phonological categories represented by elements are mapped directly onto these properties. According to this view, elements are mental constructs in the internalized grammar that manifest themselves physically as acoustic patterns. A similar view is expressed in Harris and Lindsey (2000), who describe elements as ‘auditory images’. In their view, an element is primarily a mental image of some linguistically significant information and secondarily a physical pattern in the speech signal which listeners use to cue that mental image. This makes elements crucially different from features, which are defined in terms of articulation, as in SPE, in terms of raw acoustics, as in Clements and Hertz (1991), or in terms of co-existing articulatory and acoustic specifications, as in Flemming (1995).

3.2 Element Content and Structure

Recent versions of ET assume a total of six elements, viz. |A, I, U, H, N/L, ?| (see, for example, Botma, 2004; Nasukawa, 2005; Nasukawa and Backley, 2008; for earlier proposals, see Charette and Göksel, 1994; Kula and Marten, 1998; Kula, 1999). These approaches make a distinction between what we will refer to here as ‘place’ and ‘manner’ elements.² For the purposes of this chapter, we assume that the acoustic correlates of place elements in (1a) are as in Harris and Lindsey (1995), among others. The acoustic correlates of manner elements in (1b) are as in Botma (2004).

(1) a. Place elements

<i>Element</i>	<i>Typical acoustic correlate</i>
A	central spectral energy mass (convergence of F1 and F2)
I	low F1 with high spectral peak (convergence of F2 and F3)
U	low spectral peak (convergence of F1 and F2)

b. Manner elements

<i>Element</i>	<i>Typical acoustic correlate</i>
ʔ	abrupt and sustained drop in overall amplitude
L	periodicity
H	aperiodicity

One factor that determines the phonetic manifestation of elements is their position in prosodic structure. In nuclear positions |A, I, U| have the correlates in (1a) while in non-nuclear positions |A, I, U| contribute consonantal place properties. Following, for example, Smith (1988) we assume that in such positions |A| represents velar, |I| coronal and |U| labial place.³ We consider the interpretation of manner elements in Section 5.

Another factor that determines the phonetic manifestation of an element is the occurrence, if any, of other elements in the same expression. Unlike features such as [+high] or [-sonorant], elements are phonetically interpretable without the support of other elements; Harris and Lindsey (1995) refer to this as the 'autonomous interpretation hypothesis'. In nuclear position, for example, |A|, |I| and |U| are realized as low [a], front [i] and back rounded [u], respectively. However, most segments in fact consist of a combination of elements. Such 'compound expressions' result in a signal that contains multiple acoustic patterns. The compound expression |I, A|, for example, is interpreted phonetically as /e/, its acoustic profile involving a combination of |A| and |I|.

Depending on the number of contrasts, some languages may add a further level of complexity to the way in which elements combine. The ET formalization of such combinations is based on the notion of dependency as used in DP. Dependency is relevant, for example, in languages which have a two-way contrast in the mid vowel region. Consider a language that contrasts /e/ and /ɛ/, which both involve the elements |I| and |A|. In this case it is assumed that |I| and |A| enter into a dependency relation, such that when |A| is dependent on |I| (or alternatively, when |I| is the 'head') the expression is phonetically mapped onto a close mid vowel, that is, /e/, since this vowel approximates more closely the signal pattern for |I| than for |A|. In contrast, when |A| is dominant over |I| the expression is interpreted as a low mid vowel, that is, /ɛ/. Thus, /e/ and /ɛ/ are represented as |I, A| and |I, A|, respectively, where the head of each of the expressions is underlined.

A visual analogy may help to capture the notion of dependency as used in ET. Consider two primitive categories, 'black' (2a) and 'white' (2b), as illustrated in Figure 3.3. The combination of these yields a third, complex category, 'grey' (2c).

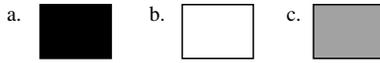


Figure 3.3

In (2c) 'black' and 'white' have equal prominence. However, suppose now that one of the two primitives can be more prominent than the other. This yields two additional categories as shown in Figure 3.4, 'dark grey' (3a) and 'light grey' (3b).



Figure 3.4

In (3ab) we can say that one of the two primitives is the head and the other the dependent, and that the two enter into a head-dependency relation. The content of the head is more prominent than that of the dependent.

With this background, we now consider some phonological arguments in favour of the representation of mid vowels along the lines suggested above. One type of support comes from processes of monophthongization and diphthongization; or as Schane (1995) calls these, 'fusion' and 'fission'. An example of the former concerns the development of early Middle English /ai, au/ to late Middle English /ɛ, ɔ:/ (cf. Anderson and Ewen, 1987: 129).

- | | | | |
|--------------------------------|---|----------------------------|--------|
| (4) Early Middle English /dai/ | > | Late Middle English /dɛ:/ | 'day' |
| Early Middle English /klau/ | > | Late Middle English /klo:/ | 'claw' |

This receives a straightforward interpretation if the elements specifying the diphthongs fuse to form a single segment with a complex element representation, with the number of timing positions left intact. Notice that the resulting monophthongs are low mid rather than high mid. Anderson and Ewen observe that this is expected, given that the original diphthongs were rising and thus had a more prominent |A| element. An example of fission is found in Faroese, where /y/ developed to /iʏ/ (cf. Schane, 1995: 592). The input to this process is a long high front rounded vowel consisting of a combination of |I| and |U|. Fission results in a structure where these elements are sequentially ordered such that the phonetic result is a short diphthong. The fact that the first element

of this diphthong is /i/ suggests that the original monophthong was headed by |I|, that is, |I, U|, representing a front high vowel with added rounding.

Further support for the relative complexity of mid vowels comes from Russian, a language with a five-vowel system /i, e, a, o, u/ and an additional length contrast. The Russian data in (5) show that in unstressed syllables /e/ neutralizes to /i/ and /o/ to /a/ (cf. van de Weijer, 1996: 9).

(5) a.	lies	[<u>l</u> 'Pɛs]	'forest'	liesa	[l'i'sa]	'forests'
	dielo	[d' <u>i</u> ɛla]	'thing'	diela	[d'i'la]	'things'
b.	gorod	[<u>g</u> 'orət]	'city'	goroda	[gəra'da]	'cities'
	god	[<u>g</u> 'ot]	'year'	v gadu	[v gə'du]	'in the year'

Van de Weijer notes that this process can be expressed in terms of the deletion of |A| and |U|, respectively, that is, /e/ |I, A| → /i/ |I| and /o/ |U, A| → /a/ |A|. This approach appropriately relates the loss of complexity with unstressed positions, capturing the observation that prosodically weak positions disfavour lexical contrasts. In a traditional feature-based account, on the other hand, there is no direct way in which this relationship can be expressed. In such an account neutralization of /e/ would involve changing [-high, -low, +back] to [+high, -low, -back]. However, notice that there is nothing intrinsically simple about [+high] as compared to [-high]; it is only in the combination with [-low] that [+high] can be said to be less complex than [-high].

As was already noted, the notion of dependency becomes relevant when we consider languages which distinguish between high mid and low mid vowels, for example, a seven-vowel system such as /i, e, ε, a, ɔ, o, u/. An approach in terms of binary [high] and [low] cannot describe such a system without the help of additional features. A dependency-based approach, on the other hand, can express such contrasts in terms of dependency relations. Thus, as noted, /e/ can be represented as |I, A|, signalling a more prominent front gesture and a less prominent low gesture, and /ε/ as (|I, A|), signalling a more prominent low gesture. In a similar vein, /o/ (|U, A|) is distinguished from /ɔ/ (|U, A|) by the former having head |U| and dependent |A|, and the latter head |A| and dependent |U|.

The preceding discussion shows that the interpretation of the notion of dependency in DP and ET differs from that in FG. In both approaches, the presence of a dependent element implies the presence of a head element. However, in DP and ET there is no implication as regards the *content* of the components that are involved (cf. also Ewen, 1995). In a compound expression |I, A|, for example, |A| is present by virtue of |I| because a dependent requires a head, not because the content of |A| requires it to be a dependent of |I|. The reason for this is that the phonetic interpretation of elements in this approach is relatively abstract, depending as it does on their position in the phonological

structure. The fact that in DP and ET the structure of representations can be motivated independently from their content makes these frameworks, in principle at least, more restrictive than FG.

3.3 A Unified Approach to Voicing and Nasality

A Number of researchers working in ET have proposed a single element representing voicing and nasality. This element is referred to either as |N| (e.g. Nasukawa, 1995, 1998, 2005) or |L| (e.g. Ploch, 1999; Kula and Marten, 1998; Kula, 1999, 2002; Botma, 2004, 2009; Botma and Smith, 2006, 2007). The difference is purely terminological; below, we refer to the element in question as |L|. An assumption that is shared by the sources cited above is that both voiced stops and nasals involve a combination of the elements |L| and |ʔ| (with |ʔ| denoting a drop in amplitude, the articulatory correlate of which is oral occlusion). However, views differ as to the specific representation of these sounds.

The first evidence that voicing and nasality are two instantiations of one and the same prime is provided by Nasukawa (1995, 1998, 2005), who offers an integrated approach to the paradoxical behaviour of nasals in Japanese. On the one hand, nasals appear to be specified for voice since they trigger post-nasal voicing assimilation (6a). Notice also that Japanese morpheme-internal NC clusters are obligatorily voiced, at least in the native Yamato lexicon (6b) (cf. Nasukawa, 2005: 3).

- | | | | | |
|--------|-----------|---------------|---|-----------------|
| (6) a. | /ʃin-te/ | [ʃinde] | | ‘die-gerundive’ |
| | /kam-ta/ | [kanda] | | ‘chew-past’ |
| b. | [ʃombori] | ‘discouraged’ | * | [ʃompori] |
| | [kaŋgae] | ‘thought’ | * | [kaŋkae] |

On the other hand, nasals do not pattern as voiced with respect to Lyman’s Law, a co-occurrence restriction on multiple voiced segments in roots. Compare the forms in (7a) with those in (7b) (cf. Nasukawa, 2005: 5).

- | | | | | | | |
|--------|----------|--------------|-------------|----|------------|---------|
| (7) a. | [sabi] | ‘rust’ | (*[zabi]) | b. | [beni] | ‘rouge’ |
| | [kazari] | ‘decoration’ | (*[gazari]) | | [nokogiri] | ‘saw’ |

Nasukawa argues that these facts receive a uniform analysis if, in the case of post-nasal voicing, |L| surfaces as a property of both parts of the NC cluster, but functions as head only in the voiced stop part. The interpretation of |L|, then, depends on its structural position: head |L| is realized as voicing,

dependent |L| as nasality. The transparency of nasals to Lyman’s Law follows directly from this: the relevant restriction can be formalized as targeting only those |L|s that function as heads.

An additional argument for the unified treatment of voicing and nasality comes from Meinhof’s Law in Bantu (e.g. Schadeberg, 1987; Kula, 2002, 2006). This concerns a process of voicing dissimilation in which a voiced stop in an NC cluster becomes a nasal if that cluster is followed by another NC cluster containing a voiced stop, as in the Luganda data in (8).

- (8) /N-genda/ [ŋŋenda] ‘I go’
 /N-bu:mba/ [mmu:mba] ‘I mould’

Kula (2002: 32), whose representation of voiced stops and nasals is similar to Nasukawa’s, accounts for this process in terms of demotion of |L| from head to dependent element, that is, |?, L| → |ʔ, L|.

In the approach of Botma (2004), which is explicitly modelled on DP representations, |L| represents not just voicing and nasalization but also sonorancy. This approach makes a distinction between place and manner elements, as outlined in (1) above, and further assumes a distinction between the manner and phonation properties of segments, which are both expressed by the elements |?, L, H|. The manner properties function as the head of the segmental structure, dominating place properties; the phonation properties are dominated by a subsyllabic constituent, for example, an onset or a nucleus. This is shown in Figure 3.5, where ‘O’ and ‘N’ are short for onset and nucleus, and ‘x’ represents the skeletal level. All structure is maximally binary branching and involves head-dependency relations, with the head-dependency relation represented by a vertical line.

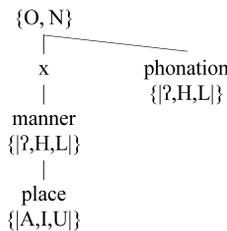


Figure 3.5

In Figure 3.5, manner and place form what may be termed the segmental ‘core’. The phonation component forms a dependent of this core. This captures the observation that it is unmarked for a segment to be specified for manner and place but marked to be specified for phonation. Notice also that following such work as Kehrein (2002) and Kehrein and Golston (2004), phonation is dominated by a subsyllabic position rather than by individual segments, as is traditionally

assumed. An advantage of this is that it captures the observation that subsyllabic constituents such as onsets are restricted to a single laryngeal contrast, which, as Kehrein shows, is borne out by cross-linguistic evidence.

In Botma's approach, |L| is interpreted as sonorancy when it occurs in the manner component, as in Figure 3.6a, the representation of back rounded /u/. The manner component of /u/ consists of |L|, which is dominated by the nucleus and itself dominates |U|. If |L| occurs in the phonation component, then its interpretation is variable. It is realized as nasalization if there is also an |L| present in the head, as in Figure 3.6b, the representation of nasalized /ũ/. It is realized as voicing if there is no |L| in the head. Figure 3.6d thus represents /b/, the voiced counterpart of the plain labial stop /p/ in Figure 3.6c (though we will argue in Section 5 that the representation of voiced stops is in fact language-specific, and may also involve the presence of |L| in the manner rather than in the phonation component).

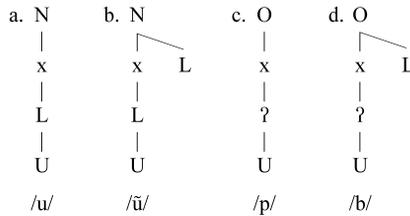


Figure 3.6

Support for this approach comes among other things from processes which trigger either voicing or nasalization, depending on whether the target is an obstruent or a sonorant. For example, in Navajo the perfective is signalled by voicing of stem-final fricatives (9a) and by nasalization of stem-final vowels (9b) (data from Rice, 1993, accent marking omitted; see also Botma and Smith, 2007).

(9) a.	Imperf	Perf	
	-ʔaaʔ	-ʔaal	'chew, eat'
	-ʔaaʃ	-ʔaaʒ	'few go'
	-loos	-looz	'lead'
b.	Imperf	Perf	
	-bi	-bĩ	'swim'
	-ʔa	-ʔã	'classificatory small object'
	-ka	-kã	'classificatory contained object'

These surface manifestations can be accounted for if the perfective morpheme is analyzed as |L|, which links to the dependent position of the stem-final segment.

In what follows, we adopt the main tenets of this approach, but diverge from Botma (2004) in two respects. The first concerns our interpretation of prenasalization, which, for Zoque at least, we take to be the phonetic realization of what is phonologically a nasalized obstruent stop. The second concerns our representation of the voiced stops of Zoque, which we believe are more appropriately analyzed as containing a specification for |L| in their manner component, contrary to structures of the kind in Figure 3.6d.

4. Voicing and Nasality in Zoque

4.1 The Zoque Data

The interaction between voicing and nasalization as found in Zoque provides an interesting test case for the idea that these two aspects are manifestations of a single phonological element |L|. At first glance, the Zoque facts appear at odds with this idea. However, we will see that they can be accounted for if we assume that voicing and nasalization operate in different prosodic domains: the former is a subsegmental property and as such part of the manner representation of segments, while the latter is a prosodic property that is dominated by onsets and nuclei.

Our focus is on the Copainalá Zoque dialect as described by Wonderly (1951a, b, c and d) and Herrera (1995) (which we refer to simply as ‘Zoque’ below). Zoque has the vowels in (10), with the quality of [i] ‘varying from mid-back to high-back’ (Wonderly, 1951b: 108). (Wonderly transcribes this vowel as [Λ], but Herrera’s transcription, which we follow here, seems more in line with its phonetic characteristics.)

- (10) i i u
 e a o

The consonants of Zoque are given in (11) (cf. Wonderly, 1951b: 105, adapted to IPA); /r, r, l/ occur in Spanish loans only, as do all non-derived instances of voiced stops (i.e. those which are not the result of post-nasal voicing assimilation).

- (11) p t tʰ ts tʃ k ?
 b d dʰ ɬ ɕ g
 f s ʃ h
 m n ɲ ŋ
 l r r
 w j

A few comments are in order regarding the phonetic realization of some of these sounds. Wonderly observes that the voiceless stops have variable aspiration, most prominently in utterance-final position. This should not be taken to mean that aspiration is contrastive, however. Rather, what seems to be the case is that final voiceless stops are necessarily released. We interpret this to mean that the sounds in question are unspecified for laryngeal structure. As for the voiced stops, Wonderly (1951b: 105) notes that these are ‘strongly articulated after nasals . . . , less strongly articulated in other positions . . . and have a spirantal character when they occur intervocally’. Voiced stops do not occur word-finally, neither underlyingly nor at the surface. On the basis of this description it seems reasonable to treat the Zoque stop system as involving a voicing contrast, with the voiced series specified in terms of an additional element |L|.

We now turn to a description of the phonological phenomena in Zoque that involve voicing and nasalization. This concerns first of all a process of post-nasal voicing, which operates regularly across morpheme boundaries: root-final and suffix-final nasals trigger voicing of a suffix-initial stop (12a and b) and prefix-final nasals trigger voicing of a root-initial stop (12c). The prefix in (12c) is the first-person singular possessive marker, which following Wonderly we represent as ‘/N-’.

(12) a.	/tukun-tan/	[tukundan]	‘earring-pl’
	/kiʔm-pa/	[kiʔmba]	‘climb-imperf’
	/min-pa/	[minba]	‘he comes’
b.	/N-ken-tam-keʔt-u/	[ngendangkeʔtu]	‘you-pl also saw it’
	/min-u-aʔm-tih/	[minuaʔndih]	‘we already came’
c.	/N-poki/	[mboki]	‘my knee’
	/N-tihk/	[ndihk]	‘my house’

Within morphemes, stops are generally voiced after nasals (13a). Wonderly (1951b: 112) notes that there are some exceptions to this, such as those in (13b), but these appear to be limited to Spanish loans.

(13) a.	[maŋ-ba]	‘he goes’
	[tsaŋga]	‘fat’
b.	[kwanto]	‘how much’
	[raŋtʃo]	‘ranch’
	[ɲempe]	‘-ever’ (‘possibly a Spanish loan’, cf. Wonderly, 1951b: 112)

Herrera analyzes forms of the kind in (13a) as having underlyingly voiceless stops. This provides a unified account of (12a and b) and (13a) in terms of

post-nasal voicing. The forms in (13b) would then have to be treated as exceptions to this process; perhaps they are lexically marked as loans. In any case, the data appear to suggest that Zoque had across-the-board post-nasal voicing prior to the influence of Spanish on the language.

The next process that we consider is prefixation of /j-/ and, more importantly for our purposes, of /N-/. We will do so against the backdrop of the analysis of these processes by Hall (2000), whose data are also drawn from Wonderly. As Hall notes, certain Zoque prefixes trigger what appears to be coalescence of the prefix and the root-initial consonant (Zoque does not have any vowel-initial roots). One such prefix is the third person singular possessive marker /j-/, as in (14).

(14) /j-pata/	[p ^j ata]	‘his mat’
/j-gaju/	[g ^j aju]	‘his rooster’
/j-faha/	[f ^j aha]	‘his belt’
/j-ʔatsi/	[ʔ ^j atsi]	‘his older brother’

/j-/ associates to any root-initial consonant, with a complication involving root-initial stridents that need not concern us here (see Hall, 2000: 725). The same process applies when a root-final /j/ is combined with a suffix-initial consonant, for example, /poj-pa/ [pop^ja] ‘he runs’. Wonderly interprets the process in (14) as involving metathesis of the prefix and the root-initial consonant. Hall (2000: 718) argues that it is more properly viewed as coalescence, that is, as involving fusion of the prefix and the root-initial consonant to a single, complex segment. Hall extends this account to include prefixation of /N-/ to root-initial stops and affricates, which he transcribes as in (15) (e.g. Hall, 2000: 720).

(15) /N-pama/	[^m bama]	‘my knee’
/N-buro/	[^m buro]	‘my donkey’
/N-tatah/	[ⁿ datah]	‘my father’
/N-disko/	[ⁿ disko]	‘my phonograph record’
/N-kaju/	[^ŋ gaju]	‘my horse’
/N-gaju/	[^ŋ gaju]	‘my rooster’

Both voiceless and voiced stops emerge as voiced and prenasalized. Apparently, then, prefixation of /N-/ neutralizes the contrast between initial voiceless and voiced stops.

While Wonderly does not comment on the segmental status of the nasal portion in initial nasal-stop sequences, Hall explicitly treats these as prenasalized stops because (1) the sequences are always homorganic, and (2) they are voiced throughout. However, since Zoque displays nasal place assimilation at prefix-root junctures (12c) and across-the-board voicing at morpheme boundaries, the

data seem equally amenable to a cluster analysis. While we agree with Hall's analysis, we believe that stronger support for a contour analysis of the NC sequences in (15) comes from the overall pattern of /N-/ prefixation.

Root-initial consonants other than stops display different behaviour under prefixation of /N-/. First of all, the prefix does not surface if the root begins with a fricative (16a) or liquid (16b) (recall that the latter occur in Spanish loans only).

- | | | | |
|---------|------------|----------|------------|
| (16) a. | /N-sik/ | [sik] | 'my beans' |
| | /N-ʃapun/ | [ʃapun] | 'my soap' |
| | /N-faha/ | [faha] | 'my belt' |
| b. | /N-lawus/ | [lawus] | 'my nail' |
| | /N-rantʃo/ | [rantʃo] | 'my ranch' |

In feature theory, this pattern has been accounted for in terms of a co-occurrence restriction on the feature combination [+nasal, +continuant] (Padgett, 1995). As far as fricatives go, this restriction appears to be violated by only a handful of Spanish loans with morpheme-internal [mf], [ns] (cf. Wonderly, 1951b: 112). However, Padgett ignores the effect of /N-/ on roots beginning with any of /j, w, h/, where most feature theories would analyze at least /h/ as also being [+continuant]. As for such roots, Wonderly (1951b: 107) observes that 'word-initial clusters ny, nw, nh are set up in which n is indeterminate as to point of articulation and is actualized as a nasalization of the y, w, h'. This is also the position taken by Piggott (1997: 467), who, based on Wonderly's characterization, provides the examples in (17):

- | | | | |
|------|-----------|----------|--------------|
| (17) | /N-jomo/ | [j̃omo] | 'my wife' |
| | /N-wakas/ | [w̃akas] | 'my cow' |
| | /N-hajah/ | [h̃ajah] | 'my husband' |

The available descriptions of this process in the literature in fact differ as to the propagation of nasalization. Wonderly suggests that only the first sound in such forms is nasalized. This is also the position in Dell (1985: 121–122), though Dell notes that some *phonetic* nasalization is likely to persist in the following vowel:

Il y a gros à parier que l'abaissement du voile du palais ne cesse pas dès la fin du glide nasal, mais se maintient durant l'émission de la voyelle suivante . . . Mais il s'agit là d'un ajustement phonétique qui ne doit pas figurer dan les représentations phonétiques [*sic*; no doubt 'représentations *phonologiques*' is intended], car il est universel. Ladefoged (1971: 33) note en

effet qu'il n'existe aucune langue où la position du voile du palais soit réglée indépendamment pour une voyelle et pour un glide qui appartient à la même syllabe que cette voyelle.

Notice that there is a discrepancy between what Dell considers a 'universal phonetic adjustment', viz. the nasalization of a vowel following a nasalized glide, and the generalization, attributed to Ladefoged, that nasalization in such sequences is a property of the syllable as a whole. Syllables are phonological constructs for which no consistent phonetic correlate has been found. In any case, it is doubtful that the syllable is relevant in light of the more recent data provided by Herrera, which show that nasalization in forms of the kind in (17) spreads rightward into the root, targeting vowels, laryngeals and glides, until it is blocked by a stop or fricative.

(18) /N-waje/	[w̃ājē]	'my mass'
/N-wʷtskuj/	[w̃ʷtskuj]	'my comb'
/N-juwi/	[jūwī]	'my skin rash'
/N-johskuj/	[jōhskuj]	'my job'

Herrera transcribes laryngeals as nasalized whenever these occur next to a nasalized vowel, despite the fact that the acoustic manifestation of nasality is likely to be greatly attenuated for [h], and absent for [ʔ]. This should not be taken to mean that the velum is raised during the realization of these sounds, of course (compare the SPE definition of [+nasal]). Indeed, forms of the kind in (19) strongly suggest that when /h, ʔ/ are followed by a vowel, the velum remains lowered.

(19) /N-ʔane/	[ʔ̃ānē]	'my tortilla'
/N-moʔot/	[mōʔ̃ōt]	'my father-in-law'
/N-hajah/	[h̃ājāh]	'my husband'

Finally, notice that the form [ʔ̃ānē] in (19) shows an interesting complication regarding the behaviour of nasals, which are apparently transparent to the nasalization induced by /N-/. Additional examples of this phenomenon are given in (20).

(20) /N-jamohk/	[j̃āmōh̃k]	'my corn cob'
/N-weʔni/	[w̃ēʔnī]	'my wasp'
/N-matsa/	[m̃ātsa]	'my star'
/N-niwi/	[ñīwī]	'my chilli'
/N-niji/	[ñījī]	'my name'
/N-jomo/	[j̃ōmō]	'my woman'

Thus, Zoque nasals do not block nasalization, but at the same time they do not trigger any nasalization themselves; neither Wonderly nor Herrera notes the occurrence of any morpheme-internal nasalization in Zoque.

4.2 Summary and Outlook

Before embarking on our analysis we first give a brief summary in the form of questions raised by the Zoque data:

- *Why do nasals trigger voicing but not nasalization?*
This requires an explanation if voicing and nasalization are both expressed by |L|. Our answer will be that voicing is active in a different domain than nasalization: the former involves an |L| that is part of the manner specification of voiced stops and nasals, the latter an |L| which is dominated by an onset or nucleus.
- *Why is nasalization triggered by the N-prefix but not by nasals?*
This follows from the assumption that the |L| in the manner component of the nasal occupies a different structural position than the |L| of the prefix /N-/.
- *What is the underlying structure of /N-/?*
We will argue that the various surface realizations of this prefix are best accounted for if the prefix consists solely of |L|.
- *What regulates the distribution of /N-/?*
/N-/ minimally associates to a root-initial segment (which is unsurprising given that /N-/ is a prefix). Association to a root-initial segment is blocked if this results in an incompatible combination of elements; association to subsequent segments is blocked for the same reason (though the relevant conditions are different).

5. Zoque Analysis

We begin our analysis by first making explicit some theoretical assumptions regarding the internal structure of segments. In this respect we follow in broad lines the approach of Botma (2004) and Nasukawa (2005). We have seen that the crucial insight of this approach is that complex expressions involve head-dependency relations, with headed elements having an acoustically more prominent signature than dependent ones. The representations that we propose here are based on those in Botma (2004), though we depart from this

approach in a number of respects. One important difference is that we assume here that the voicing of voiced stops (in Zoque at least) is represented by a dependent |L| in the manner component and not, as in Figure 3.5, in the phonation component.

We assume, first of all, that plain oral stops in Zoque are specified for |ʔ| in combination with a place element, as in Figure 3.7a), the representation of /p/. Voiced stops differ from voiceless ones in having an additional dependent |L|, as in Figure 3.7b, the representation of /b/. Voiceless fricatives, such as /f/ in Figure 3.7c, differ from voiceless stops in having an additional element |H| in head position, signalling a relatively prominent noise component (i.e. oral friction) and a relatively less prominent drop in amplitude (a property of continuants). Affricates, such as /ts/ in Figure 3.7d), have the reverse dependency relation, signalling an acoustically prominent drop in overall amplitude (a property of stops) and a relatively less prominent noise component (marked by their fricated release). Notice also that we assume that affricates consist of a single manner component, so that the stop and fricative portion are unordered phonologically (in Section 5.1 we will propose the same type of representation for the prenasalized stops of Zoque). Finally, we assume that nasals are like voiced stops, but with the reversed dependency relation holding between |ʔ| and |L|, signalling a greater degree of periodicity. Thus, the representation of /m/ is as in Figure 3.7e.

a. O	b. O	c. O	d. O	e. O
x	x	x	x	x
ʔ	ʔ	H	ʔ	L
U	L	ʔ	H	ʔ
	U	U	I	U
/p/	/b/	/f/	/ts/	/m/

Figure 3.7

Notice in Figure 3.7e that we assume that nasals have |L|, that is, ‘periodicity’, as an integral part of their manner component. This implies that we do not recognize a phonological category of voiceless nasals. Indeed, cross-linguistic evidence suggests that nasals which qualify as voiceless phonetically are more appropriately analyzed as aspirated phonologically (see, for example, Ewen, 1980; Lombardi, 1991a). In our approach such nasals contain an |H| in the phonation component, similar to aspirated stops (see Botma, 2003).

A comment is also in order regarding the phonetic exponents of Figure 3.7e. In most languages, including Zoque, this structure is realized as a nasal consonant, presumably because the acoustic signature of nasals offers perceptually the most salient compromise between ‘stopness’ (i.e. |ʔ|) and relatively prominent ‘periodicity’ (i.e. |L|). However, the structure can also be realized as a voiced oral stop. Botma and Smith (2006, 2007) claim that this is the case, for example, in languages where voiced oral stops alternate with nasals under nasal harmony, such as Cama (Ebrié) and Southern Barasano. In Botma and Smith’s analysis these languages have underlying voiced stops, which surface as nasals through association of dependent |L| in their phonation component, in much the same way as nasalization operates in Zoque. The variable interpretation of the structure in Figure 3.7e is important since it shows that our approach has no ‘double specification’ of nasality. Rather, nasality is one possible phonetic interpretation of a manner component in which |L| dominates |ʔ|, while, as we will see below, nasalization is uniformly represented by an |L| specification at the prosodic level.

5.1 The /N-/Prefix and Nasalization

We suggest that the /N-/ prefix consists of an |L| element, which is minimally realized as a property of the root-initial consonant (recall that Zoque lacks vowel-initial roots).⁴ Consider first the realization of /N-/ in roots with initial /j, w/. Although the effect of the prefix on such forms is the most dramatic, the analysis that we suggest for this pattern is straightforward: we propose that /N-/ links to the phonation component of the root-initial consonant, from which it spreads rightwards to all eligible targets. This is shown in Figure 3.8 for the form /N-jomo/ ‘my woman’ (cf. (20)), where dotted association lines indicate spreading of |L| to each nasalizable target.

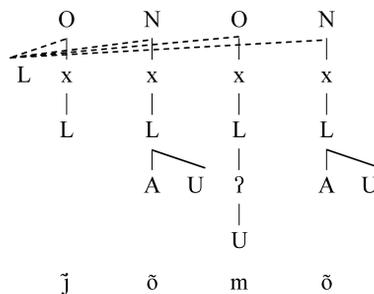


Figure 3.8

All segments in Figure 3.8 are suitable nasalization targets. We suggest that this is due to the fact that they are all sonorants, and are identified as such by their having a manner component headed by |L|. To this extent, the pattern of nasalization in Zoque is like that of Amazonian languages like Tuyuca and Southern Barasano, where nasal harmony targets all sonorants in the harmonic domain (see, for example, Piggott and van der Hulst, 1997; Botma, 2009). Segments not headed by |L| either reject nasalization (fricatives and liquids) or accept it but fail to pass it on (stops), for reasons discussed shortly.

Notice in Figure 3.8 that although /m/ is targeted by nasalization, association of |L| here does not have an observable phonetic effect. The reason for this is that the combination of head |L| and dependent |ʔ| in the manner component already identifies the segment as a nasal. On the assumption that nasalization operates exclusively at the subsyllabic level, that is, at the level of onsets and nuclei, this provides a straightforward explanation for the fact that nasals do not trigger any nasalization themselves: their lexical representation lacks an |L| specification at the appropriate level in the phonological structure.

A comment is also in order regarding the behaviour of /h, ʔ/, which, as we have seen, are targets for nasalization. One way to account for this is to assume that the range of nasalization targets includes placeless segments, in line with the approach in Botma (2004). Another would be to analyze /h, ʔ/ as glides, in which case they would contain a head |L| in their manner component.

5.2 Prenasalization

We now consider those cases in which /N-/ associates to a root-initial stop. We have seen that such stops surface as voiced and prenasalized. Our analysis of this is shown in Figure 3.9, for the form /N-pama/ ‘my knee’ (cf. (15)).

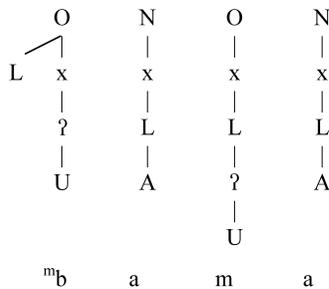


Figure 3.9

Our claim, then, is that obstruent stops also provide a target for nasalization. The result of this is a nasalized obstruent stop, which is realized phonetically as prenasalized. The advantage of this is that it provides a uniform interpretation of /N-/ prefixation in terms of the association of |L| to (minimally) the root-initial consonant – though notice that it also entails that the range of nasalizable sounds in Zoque is not restricted to the class of sonorants alone, but also includes obstruent stops, at least in root-initial position.

Our interpretation of Zoque prenasalization raises three questions. The first concerns the observation that in Figure 3.9 the oral stop portion and the nasal portion are unordered, similar to what we suggested for affricates above. This is contrary to the account of, for example, Anderson (1976), who claims that prenasalized stops display ‘edge effects’ and so must involve ordering of the oral and nasal portions. We do not believe that such ordering is always necessary, however. There does not seem to be any compelling phonological evidence for a contour representation of prenasalized stops in Zoque. We therefore suggest that prenasalization in this language is a result of phonetic implementation. The fact that nasalized stops are realized as pre- rather than as post-nasalized has a reasonable perceptual motivation, in view of the more salient release cues of the former.

Another question concerns the question why nasalized stops are voiced, despite the fact that they lack an |L| in their manner component. We suggest that the answer to this is that single segments cannot display internal voicing contrasts. The more specific question is then why prenasalized stops are realized as fully voiced rather than voiceless. We suggest that the answer to this is that the presence of phonological nasalization implies the presence of phonetic voicing – that is, we know of no instances where nasalization is an exclusive property of voiceless segments.⁵ If such segments are indeed unattested, then we can say that any |L|, irrespective of its position in the phonological structure, signals the presence of phonetic voicing, including in those cases where its phonological function is to encode nasality, as in Figure 3.9.

Our interpretation of prenasalized stops as nasalized obstruent stops requires a more detailed discussion than is possible here. However, we believe that our perspective is supported by cross-linguistic evidence. In a recent typological study, Riehl and Cohn (2010) observe that languages never seem to contrast pre- with post-nasalized stops, which suggests that the ordering of the stop and nasal portions need not be specified in the phonological representation of these sounds. Riehl and Cohn further reject a category of voiceless prenasalized stops, which suggests that prenasalized stops need not have a phonological specification for voicing.

The final question that must be considered is why in Figure 3.9 nasalization is limited to the root-initial stop. We suggest that this is due to a general

sharing restriction on ‘floating’ |L|. This |L| can be a property of either an obstruent stop, in which case it is associated exclusively to the root-initial onset, as in Figure 3.9, or of sonorants, in which case it is minimally associated to a root-initial consonant and a following vowel, as in the forms in (18), (19) and (20). Thus, representations of the kind where |L| is simultaneously linked to an obstruent and a sonorant, as shown in Figure 3.10, are ruled out.⁶

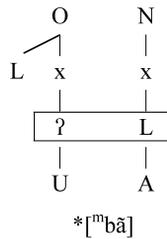


Figure 3.10

In traditional autosegmental terms, the ungrammaticality of [‘mbã] is accounted for by specifying the prenasalized stop in terms of both [+nasal] and [-nasal], which, since they are ordered on the same tier, rules out spreading of [+nasal] to the following vowel (Goldsmith, 1979b; Anderson, 1976). The sharing restriction in Figure 3.10 accounts for the same observation without recourse to contour representations or binary features. Notice that it also accounts for the observation that a root-initial stop in Zoque is simultaneously a target and a blocker of nasalization. We suggest that Figure 3.10 is also relevant in those languages in which nasal harmony targets sonorants but not obstruents; for examples of such languages, see, for example, Piggott (1992) and Piggott and van der Hulst (1997).

5.3 Absence of the /N-/ Prefix

We next consider those instances where /N-/ fails to surface, viz. in the case of root-initial fricatives and liquids (cf. (16)). We suggest that this is due to another co-occurrence restriction, viz. one which bans the combination of ‘floating’ |L| and |H|.

The ban on nasalized fricatives implies that /N-/ remains unrealized, given that it cannot be left-aligned with the root (i.e. *[fãhã]) as illustrated in Figure 3.11.

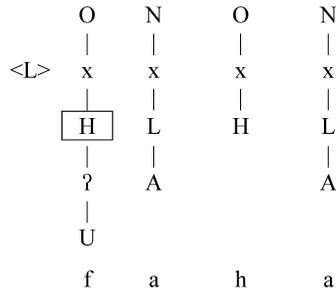


Figure 3.11

This restriction on |H| and |L| provides a straightforward account of the behaviour of fricative-initial roots.⁷ However, it seems reasonable to assume that the behaviour of root-initial liquids requires a more detailed account since, in the approach of Botma (2004) at least, these contain |H| and |L| as part of their manner specification. Reasons of space do not permit us to discuss this issue in any detail. However, a feasible hypothesis would be that the restriction on |H, L| combinations holds only for derived structures.

5.4 Post-Nasal Voicing

Finally, some brief comments are in order concerning our interpretation of post-nasal voicing. What is crucial for our purposes is that this process, unlike nasalization, applies at the level of segment-internal structure. Specifically, in line with Kula (1999, 2002) and Nasukawa (2005), we assume that post-nasal voicing involves the addition of |L| to the manner component of the affected stop, under the influence of the preceding nasal. This is shown in Figure 3.12 for the coda-onset cluster in /min-pa/ 'he comes' (cf. (12a)), where the manner component of the stop contains an inserted |L|.

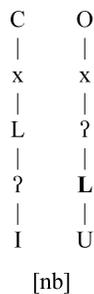


Figure 3.12

Notice that this interpretation departs from Nasukawa's in that we take nasality to be represented by a head |L| and voicing by a dependent |L|, rather than vice versa. Hence, a possible restriction on this kind of process is that inserted elements cannot function as head.

6. Conclusion

This chapter has offered an overview of a number of different theoretical approaches to phonological primes, the smallest units of phonological representation. We hope to have shown that from the point of view of theoretical parsimony, approaches which assume a reduced number of privative primes, such as ET, are, in principle at least, more restrictive than approaches which assume equipollent features, such as traditional distinctive feature theory. Clearly, the challenge facing restrictive approaches is to account for the same phonological patterns with a smaller number of primes. One such challenge, as we have seen, is the phonological behaviour of nasality and voicing in Zoque. We have attempted to show that the Zoque facts can be given a straightforward and non-stipulative account on the assumption that voicing and nasalization are represented by one and the same prime, viz. the element |L|, whose exponence depends on its position in the phonological structure. More specifically, we have argued that nasalization, which in Zoque is triggered by the /N-/ prefix, operates at the level of onsets and nuclei, while voicing, which is triggered by nasals, operates at the segmental level.

7. Acknowledgements

Bert Botma acknowledges financial support from the Netherlands Organization of Scientific Research (NWO), project no. 2008/01214/GW. Kuniya Nasukawa acknowledges financial support from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government under grants 19520429 and 22320090.

8. Notes

1. However, notice that a lowered velum does not imply the presence of nasal airflow. This issue is relevant in languages where the domain of nasalization apparently includes [ʔ], such as Zoque. For discussion of the status of nasalized laryngeals, see, for example, Walker and Pullum (1999), Botma (2004).
2. Nasukawa and Backley (2008) refer to these as 'edge' and 'resonance' elements, respectively.

3. Another approach to consonantal place can be found in Nasukawa and Backley (2008), where |A| represents the category 'guttural' and is typically interpreted as pharyngeal, |I| represents the category 'coronal' and is typically interpreted as dental, and |U| represents the category 'dorsal' and is typically interpreted as labial/velar (cf. [grave] in Jakobson and Halle, 1956, [peripheral] in Rice and Avery, 1991, [dark] in Backley and Nasukawa, 2009a).
4. In this analysis |L| is similar to what is called a 'floating feature' in autosegmental phonology.
5. We do not claim that voiceless sounds cannot be realized with nasal airflow; this has been observed for the voiceless fricatives of Coatzespan Mixtec, for example (Gerfen, 1999). However, such voiceless sounds are always contiguous to voiced ones, and our claim is therefore that there is no reason to treat them as nasalized phonologically. A reasonable option is to attribute their nasalization to the effect of phonetic interpolation, in line with the approach of Nasukawa (2005).
6. This sharing restriction reflects a weaker version of the claim that voicing and nasalization are in complementary distribution (Botma, 2004, 2009; Botma and Smith, 2006, 2007). We believe that this claim is too strong in view of the existence of nasalized stops, as in Figure 3.9. However, this does not imply that sonorants and obstruents may display different behaviour with respect to nasalization, as is shown by the impossibility of sequences like *[^mbā], and by the patterning of these classes in languages like Southern Barasano.
7. Zoque also lacks voiced fricatives, a natural representation of which would involve a manner structure in which |H| dominates |L|, parallel to the representation of voiced stops proposed in Figure 3.7b.

4 Syllables

Péter Szigetvári

Chapter Overview

1. Why Have Syllables?	65
2. Underlying or Derived?	67
3. Internal Structure	73
4. Vowelless Syllables	79
5. Syllables without Constituents	81
6. Conclusion	91
7. Acknowledgements	92
8. Notes	92

Syllables are one of the few phonological notions that the average layman knows about. Others include sounds, consonants and vowels.¹ Syllables belong to the prosodic domain of sound structure, together with feet: both categories are related to the rhythm of speech. While feet are the manifestation of the rhythmic alternation of stressed and unstressed portions, syllables represent the alternation of high-sonority (syllabic) and low-sonority (nonsyllabic) portions, or vowels and consonants, as they are usually referred to. These two waves, the rhythm wave and the sonority wave, carry the speech signal. Their interference is constructive, that is, a rhythmic peak always coincides with a sonority peak, so the two waves never cancel out.

Since, unlike morphemes, syllables are not associated with meaning, semantic considerations do not enable the speaker to discern where one syllable ends and the next one begins. In fact, it is not even obvious if phonological theory has to include an entity that corresponds to syllables. It is well known that syllables are claimed to be unnecessary by Chomsky and Halle (1968) in what came to be a cornerstone of modern phonological theory, *The Sound Pattern of English* (SPE). Section 1 shows that although technically viable, this solution misses a generalization, the recurrent environment that can be identified as the end of

the syllable, which can only be expressed disjunctively in the classical version of the generative theory. On the other hand, the syllable as a unit of phonological representations does not appear to be absolutely necessary. Section 2 examines whether syllable structure can be derived from the order of sound strings, or at least some part of it must be given lexically. The impossibility of equating word and syllable boundaries is demonstrated. This calls for an exclusively sonority based syllabification, which also turns out to be problematic. The conclusion is that syllabicity must be lexically marked.

Arguments for the possible internal organization of the syllable are looked at in Section 3. They are based on syllable weight (Section 3.1), phonotactic constraints (Section 3.2) and closed syllable shortening (Section 3.3). None of these are capable of clearly distinguishing between the CV+C and the C+VC basic split within a syllable. Section 4 introduces the option of having ‘syllables’ without a pronounced vowel, a move that reshuffles most of what has been thought of about syllables. (Interestingly, the first written record of this analysis comes from the middle of the eighteenth century.) This idea is elaborated in Section 5, showing that a model that involves only relations between pairs of segments and dispenses with the prosodic hierarchy above them has certain advantages over traditional models of syllabic structure in representing consonant clusters (Section 5.1), long vowels (Section 5.2), the status of intervocalic consonants (Section 5.3) and syllable weight (Section 5.4).

1. Why Have Syllables?

Early generative phonological descriptions avoid explicit reference to the syllable.² They deny not only the level of the syllable in the prosodic hierarchy, but also its boundaries, although representations are crowded with boundary symbols – albeit mostly grammatical ones. The SPE, for example, proposes the rule in (1) for French elision and liaison (Chomsky and Halle, 1968: 353).

(1) SPE elision and liaison rule

$$\left[\begin{array}{l} -\alpha \text{ vocalic} \\ \alpha \text{ consonantal} \end{array} \right] \rightarrow \emptyset / _ \# [\alpha \text{ consonantal}]$$

In its own clumsy way the rule aims at retaining the consonant – vowel alternation across a word boundary by deleting a consonant before a consonant and a vowel before a vowel. Note that the rule in (1) deletes a vowel, and does not delete a consonant, before a glide (e.g. *l'oiseau* [lwazo] ‘the bird’ and *petit oiseau* [pətitwazo] ‘little bird’). But since glide-initial words labelled [+foreign] by Chomsky and Halle pattern with consonant-initial words (e.g. *le yogi*, *le watt*),

the feature [syllabic] is tentatively introduced to replace the original [vocalic]. The rule in (2) takes care of elision and liaison in foreign words (1968: 355).

(2) SPE elision and liaison rule for foreign words

$$\left[\begin{array}{l} -\alpha \text{ syllabic} \\ \alpha \text{ consonantal} \end{array} \right] \rightarrow \emptyset / _ \# \left[\begin{array}{l} -\alpha \text{ syllabic} \\ + \text{ foreign} \end{array} \right]$$

We see in (2) that Chomsky and Halle are pressed to posit a feature [\pm syllabic] in a framework that otherwise denies the necessity of including this notion in its formalism.

To explain the divergent behaviour of obstruent+glide/r consonant clusters with respect to stress placement rules, the definition of a weak cluster (the SPE equivalent of a light syllable) in (3) is arrived at (1968: 83).

(3) The weak cluster in SPE

$$\left[\begin{array}{l} - \text{tense} \\ \text{V} \end{array} \right] \text{C}_0^1 \left[\begin{array}{l} \alpha \text{ vocalic} \\ \alpha \text{ consonantal} \\ -\text{anterior} \end{array} \right]_0$$

The formula in (3) includes the feature [-anterior] to exclude [l] (while including [j w r]), since in their analysis words like *armadillo*, *vanilla*, *umbrella* contain a geminate /ll/, hence a strong cluster, which as a result is stressed. However, words like *discipline*, *pánoply*, or in Chomsky and Halle's analysis *míracle*, *clávicle* etc. (1968: 197) show that stop+l sequences do create a weak cluster, since we find antepenultimate stress in these words, which is expected only if the penultimate vowel is followed by a single consonant or a weak cluster (e.g. *citizen*, *mímicry*). Thus, besides being extremely unnatural, the definition of the weak cluster is empirically inadequate.

Kahn (1976: 22ff.) notes that the environment $_ \left\{ \begin{array}{c} \text{C} \\ \# \end{array} \right\}$ is recurrent in many phonological rules across languages, it is here, for example, that [l] is velarized and [r] is deleted in Standard British English. The attempts at making consonants and the word boundary a natural class to avoid the disjunction turned out to be unsuccessful. Chomsky and Halle seem to be missing a generalization here, the syllable boundary.

Subsequent research has adopted the notion of the syllable in phonological theory collecting a set of arguments for its necessity. Phonotactic constraints are often referred to for justifying syllables. But, as Blevins admits, 'surprisingly, there are few if any feature co-occurrence constraints which appear to take the syllable as their domain' (1995: 237, fn. 8). In fact, as I will argue in Section 3.2, phonotactic constraints hold primarily between adjacent consonants (and

adjacent vowels, that is, diphthongs). Crucially, interdependent consonants may or may not be tautosyllabic, it is their adjacency that matters. A tautosyllabic onset and coda³ do not have any effect on each other, heterosyllabic coda+onset clusters, however, do. Therefore, the syllable is not a relevant domain for formulating phonotactic constraints.

Another argument for the unity of the syllable is that stress anchors not on vowels, but on whole syllables (Hayes, 1995a: 49f.). Halle and Vergnaud claim that vowels are the stress-bearing elements: in some languages stress can fall on either mora of a long vowel (1987: 49, 61, 193). If this is a possibility, the syllable is not indispensable for locating stress. But even if it were the case that only the first mora of a syllable could bear stress, one could still argue that the anchor for stress is not the syllable but the nucleus or the rhyme in it.

As has already been mentioned, native speakers have intuitions about syllables in their language. Harris and Gussmann (1998) argue that these intuitions are the result of linguistic traditions. They examine the case of word-final consonants, which, according to what they refer to as the 'western' or Graeco-Roman tradition, belong to the same syllable as the preceding segment(s). According to a more ancient 'eastern' tradition, such a consonant is the onset of a syllable that lacks an audible nucleus. Native speaker intuitions about syllables – at least some of them – are thus not based on unconscious linguistic analysis, but on traditions about language learnt at school. They should accordingly be treated with suspicion.

Before one hastily concludes that the notion of the syllable is useless, it is worthwhile clarifying what it means 'to have' syllables in one's theory. The periodic alternation of low- and high-sonority portions in sound sequences is an empirical fact of sound-based human languages (noticed by Sievers, 1881), this fact must be accounted for. If the way a particular theory accounts for the sonority wave in speech is called 'syllable', then any theory is bound to have syllables. If, however, we restrict the use of this term to an explicit node in the prosodic hierarchy, then theories will vary to a large extent in having or not having syllables. Accordingly, the only meaningful way in which two theories can be distinct in one having and the other one lacking syllables is having an explicit bit in the representation that is labelled 'syllable', or 'rhyme', or at least having a 'syllable boundary'. But even if a theory does not recognize the syllable explicitly, the term can be used in a descriptive way to denote a sonority cycle of the speech signal.

2. Underlying or Derived?

Blevins lists three reasons for assuming that syllable structure is not underlying, but derived (1995: 221): (i) the scarcity of minimal pairs distinguished only

by syllabification, (ii) syllabicity alternations (e.g. *opin*[iə]~*opin*[jə]n); and (iii) morphemes often cannot be parsed according to the syllabification rules of the language. The first two reasons are closely related, and are primarily a result of the ambiguous status of glides. This will be discussed in Section 2.3. Let us first see if reason (iii) provides a solid foundation for derived syllabification.

The ill-formedness of the syllable structure of morphemes in the lexicon is trivially true for suffixes consisting only of consonants. The English past tense suffix *-d*, for example, cannot be syllabified to create a well-formed word. This is not surprising: *it is* not a well-formed word. However, it is not only affixes but also stems that resist exhaustive syllabification. A word like *hymnal* [hɪmnəl] suggests that *hymn* [hɪm] ends in the cluster [mn] in the lexicon. Since this cluster does not occur syllable finally, it seems to be a plausible assumption that lexically it is represented as an unsyllabified sound string //hɪmn//, which surfaces as [hɪm] after it turns out that the final [n] cannot be incorporated by syllabification and hence is deleted by stray erasure (Steriade, 1982).⁴ When suffixed with *-al*, the final [n] is rescued because it becomes the onset of the second syllable, therefore it is pronounced.

Such facts argue against underlying syllabification only if we accept two further assumptions: namely, that a lexical form must be exhaustively syllabified and that the stem of *hymnal* is a lexical form. Neither assumption is self-evident. The morpheme *hymn* could lexically contain a fully syllabified part [hɪm], followed by an unsyllabified [n]. More importantly, it is not obvious that //hɪmn// is an entry of the lexicon. Kaye demonstrates that forms involving nonanalytical morphology (of which *-al* suffixation is an example) are invisible to phonology (1995: 308f.). This means that *hymnal* is a lexical entry in its entirety, creating a single phonological domain, therefore phonology is ignorant of its internal morphological structure. In the case of analytical affixation, the situation is different: the stem of *hymning* [hɪmɪŋ] does undergo phonological processing. As a result, it surfaces in the same form as the unsuffixed stem *hymn*. What this means is that there is a lexical entry //hɪm// (without a final [n]), and another entry //hɪmnəl//, the two are not related by phonological rules. According to this view, lexical entries are free forms (complete words), hence their syllabification can be lexical. As for affixes, nonanalytical ones (like *-al*) do not occur independently in the lexicon, while analytical ones (like *-ing* or *-d*) are not phonological domains in themselves, they create a domain together with the stem they attach to.⁵

If the syllabification of a string is predictable from the phonetic properties of the segments it is made up of, the question whether syllable structure is lexically given or derived is only relevant if there are phonological rules that apply lexically prior to syllabification taking place. Without lexical rules, the question loses its significance. The point of having syllabification algorithms is then to check the phonotactic validity of segment strings.

2.1 Syllabification with Reference to Word Edges

Kahn's (1976) syllabification algorithm consist of a set of rules that identify the syllabic segments in a string and then stack nonsyllabic segments to it, first to its left, then to its right (these latter rules are referred to as rule IIa and IIb, respectively). The order is important: if an intervocalic consonant (cluster) can be syllabified to both the preceding and the following vowel, the latter choice is preferred: resulting in onset maximalization. Thus, *Lisa* is split as [li:sə] (although it could be [li:s.ə]) and *supra* as [su:pra] (although it could be [su:p.rə], but *hangar* as [hæŋ.ə] and *anger* as [æŋ.gə] (here no other option is available). Kahn claims that the set of permissible syllable-initial (and syllable-final) consonant clusters is part of rules IIa and IIb (1976: 45). Actually, permissible clusters at the edges of the syllable are exactly those that we observe at the edges of the word, the reason being that words that cannot be exhaustively syllabified by rules IIa and IIb are ungrammatical. Thus, although the stacking operation of rule IIa is not controlled by reference to the set of word-initial consonant clusters but by a list which is part of the rule, the two sets are equivalent, so in effect a syllable-initial cluster is possible if that cluster occurs at the beginning of words in the given language, and the same holds for syllable- and word-final clusters.

The effect of this algorithm could be exemplified by the parsing of *anxiety*: in the classical generative model its morphological connection to *anxious* [æŋkʃəs] suggests an underlying [-ŋgz-] in *anxiety*. Since [-ŋg] does not occur syllable/word finally and [gz-] does not occur syllable/word initially, the form [æŋgzaiəti] should be unparseable and hence ungrammatical. Wells (1990), nevertheless, does list it as a possibility. Lowenstamm mentions similar problems in Finnish, where three-member internal clusters occur, but two-member clusters are not found at either edge of a word (1981: 601). Ancient Greek has -lp- within words (e.g. [elpis] 'hope'), but neither initial *#lp-, nor final *-l# or *-lp# is found, thus there seems to be no way to syllabify this word, if we want to maintain that the phonotactics of syllable and word edges is the same.

Kahn's predictions are defied at another point by Ancient Greek. His algorithm takes consonants one by one to check whether the given string is a possible cluster or not. Thus, if VC_i is a nonexistent string, VC_iC_j must also be impossible, since the algorithm has no chance to examine the longer cluster if it stops at the shorter one. Now word finally, as we have seen, *-l# is not possible, but -ls# is (e.g. [hals] 'salt'), albeit marginally. The situation is the same with *-p# versus -ps# (e.g. [gy:ps] 'vulture', [ɔ:ps] 'eye' etc.), with a larger set of examples. Longer clusters also exhibit this unexpected pattern: *-ŋk# versus -ŋks# (e.g. [laryŋks] 'larynx', [sp^hyŋks] 'sphynx').

2.2 Lowenstamm's Universal Syllable

Lowenstamm (1981) shows that Kahn's algorithm is not only empirically but also theoretically faulty, since it cannot syllabify strings that are not encountered on the surface.⁶ He claims that in [gɪdɔrim] (the plural form of Yiddish [gedır] 'limitation') the first vowel is inserted by a rule. Epenthesis thus takes the input [gdɔrim] and yields [gɪdɔrim]. Since Yiddish does not exhibit the cluster [gd-] word initially, [g] must be left unsyllabified by Kahn's rule IIa. But if it is not incorporated into any syllable, the epenthesis rule cannot apply, in fact, the whole form is discarded as ungrammatical. The rationale of the epenthesis rule is exactly to correct the ill-formed word-initial cluster, but it has no chance to do so in Kahn's framework.

Lowenstamm proposes a return to the notion of sonority in assigning segments to syllables. Following Jespersen (1920) and Grammont (1933), he defines a 'universal syllable' as in (4).⁷

(4) The Universal Syllable

In a string of segments, a syllable is a maximal substring such that

- a. (i) no segment is lower on the hierarchy than both its immediate neighbours
- (ii) no two segments of equal ranking on the hierarchy are adjacent
- b. the onset is maximal within the limits of (a).

The hierarchy referred to in (4) is, of course, the sonority hierarchy. The effect of (4 a. (i)) is that within a syllable, sonority first rises to a peak, then it falls away from the peak. Both the rising phase and the falling phase may be missing, but a fall in sonority may never be followed by a rise within any syllable. This is the sonority sequencing principle. The sonority of segments thus seems to play an indispensable role in defining possible and impossible syllables. Let us examine sonority.

2.3 Sonority

The sonority hierarchy is a rank list of segment types – in some analysts' view of individual segments (e.g. Ladefoged, 1993) – from least to most sonorous. A phonetic definition of sonority and whether the sonority of segments is a primitive of phonological theory (e.g. Kiparsky, 1979) or derived from their phonetic properties (e.g. Clements, 1990; Harris, 1990) need not concern us here.

For some languages a very crude sonority hierarchy, containing two steps, will do: $C \rightarrow V$ (where the arrow points from minimal to maximal sonority). These languages will have neither consonant nor vowel clusters (i.e. long vowels or diphthongs), but a strict alternation of consonants and vowels, that is, $\#(C)VCV \dots C(V)\#$. Some languages that allow a maximum of two consonants to occur next to each other may also be analysable with this simple scale, since in such languages (e.g. Japanese or Manam) the boundary of the syllable will always be between the consonants, as ruled by (4 a. (ii)).⁸

An interesting situation arises in systems that contain long vowels. If a syllable with a long vowel is marked as CVV, this appears to violate (4 a. (ii)): two segments of equal ranking are adjacent within a syllable.⁹ The problem is only apparent, and lies in our linear–non-linear hybrid rendering of the long vowel. In a genuine linear framework a long vowel is just another simple feature matrix (i.e. a single segment) with the feature [+long] included. In a nonlinear representation a long vowel looks like (5), where V stands for the melody of any vowel.

(5) The nonlinear representation of a long vowel



That is, the graphical representation 'VV' for a long vowel is the result of merging two tiers, the skeletal and the melodic tier. If we are to keep using sonority sequencing, a long vowel is to be taken as a single unit.

To get diphthongs under the control of (4) the sonority hierarchy has to be refined: the glide part of the diphthong has to be assigned a sonority value smaller than the vowel part.¹⁰ The resulting scale is this: $O \rightarrow S \rightarrow V$ (where O means obstruent, S means sonorant consonants; this scale is adopted by Zec, 1995).

This three-term sonority scale is enough for controlling sonority sequencing in an overwhelming majority of cases. Apart from [s], two obstruents or two sonorant consonants rarely (if ever) occur together on the same side of the vowel within a syllable (i.e. *OOV, *SSV, *VOO, *VSS), consequently obstruents or sonorant consonants do not have to be sequenced among themselves, and so they do not have to be ranked. Departures from this setting occur predominantly at word edges, but as we are going to see in Section 4, there is reason to reject the automatical tautosyllabic analysis of any word-initial or word-final string of consonants. If such a view can be supported, then words like *film*, *act*, *sky* etc. do not require ranking sonorants or obstruents in the

sonority scale. This simple scale becomes inadequate for systems where sonorant consonants can form the sonority peak in the syllable: for example, the first two consonants of Czech *mlha* 'fog' would be split by a syllable boundary, if all sonorants were of the same rank. In reality, the syllabification is [mɫ.fia], with syllabic [ɫ]. The resulting scale is the following: O→N→S→V (where N stands for nasal, S for nonnasal sonorant consonants).

Another case where the ranking of sonorants seems necessary is for words like English *wild* and *kind* where a diphthong occurs in a closed syllable. Here the glide and the liquid/nasal have to be ordered, such that the glide is higher on the scale: O→N→L→G→V (where L means liquids and G glides; this scale is advocated by Clements, 1990).¹¹

The patterns of syllabic consonant formation in English also call for the splitting of liquids and nasals, but on the other hand, it questions the previous split of glides and other sonorant consonants. Let us first see the basic pattern in (6). (The judgements are those of Wells, 1990.)

(6) Syllabic consonants in English: the basic pattern

a. little	lɪtəl	~ -tɫ	b. colonel	kɜ:nəl	~ -nɫ
people	pi:pəl	~ -pɫ	camel	kaməl	~ -mɫ
hassle	hasəl	~ -sɫ	plenary	plɪ:nəri	~ -nɾ
awful	ɔ:fəl	~ -fɫ	camera	kamərə	~ -mɾ
cotton	kɒtən	~ -tɳ	salary	saləri	~ -lɾ
special	speʃəl	~ -ʃɫ	c. melon	mɛlən	~ *-lɳ
Belgian	bɛlʤən	~ -ʤɳ	column	kɒləm	~ *-lɳ
fathom	faðəm	~ -ðɳ	canon	kanən	~ *-nɳ
Benjamin	bɛnʤəmɪn	~ -ʤɳ	minimum	mɪnɪməm	~ *-mɳ
luxury	lʌkʃəri	~ -ʃɾ	lemon	lɛməm	~ *-mɳ
peppery	pɛpəri	~ -pɾ	venom	vɛnəm	~ *-nɳ

The data in (6a) aim to illustrate that a sonorant consonant may be syllabic after any obstruent.¹² The obvious and common explanation for this fact is that obstruents are less sonorous than sonorants, accordingly the syllabic sonorant remains the sonority peak after the loss of the vowel. As (6b) and (6c) show, syllabic consonant formation is more restricted after sonorant consonants. The data suggest the following ranking within sonorants: m, n→l→r. Syllabic glides (i.e. high vowels) may occur after any sonorant consonant, that is, they should be located at the top of the sonorant subscale: m, n→l→r→i/j, u/w. The data presented in (7), however, partially reverse the ranking.

(7) Syllabic consonants in English after glides

a. barrel	barəl	~ -r̩	b. narwhal	nɑ:wəl	~ -w̩
barren	barən	~ -r̩	loyal	lɔjəl	~ -j̩
quorum	kwɔ:rəm	~ -r̩	tower	taʊwər	~ -w̩
terrorist	tɛrərɪst	~ -r̩	lawyer	lɔjər	~ -j̩

These data show that syllabic [l] and syllabic nasals behave as if they were higher up the sonority scale than [r], (7a), and the glides [j] and [w], (7b).¹³ This means that the syllabic and nonsyllabic versions of high vowels/glides occupy separate positions on the sonority scale: the nonsyllabic versions are below all sonorant consonants, the syllabic versions are above them, together with vowels. If so, then either the high glides [j] and [w] and – what invites to be called the low glide – [r] and their syllabic (i.e. vocalic) counterparts are different entities, or syllable structure cannot be derived, but is lexically given. Obviously, the syllabicity of these segments cannot be derived by reference to their sonority, if we want to derive their sonority by reference to their syllabicity. (8) depicts this lexical marking: V stands for the nucleus or peak of the syllable and C for the syllable margin.

(8) Syllabic and nonsyllabic ‘glides’

a. [i]	b. [j]	c. [u]	d. [w]	e. [ɚ]	f. [r]
V	C	V	C	V	C
i	i	u	u	r	r

(8a), (8c), and (8e) represent the versions of these segments occurring in the syllable peak, (8b), (8d), and (8f) represent their syllable marginal occurrence.

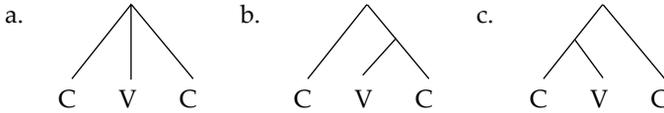
Accordingly, at least syllabicity must be lexically given, and cannot be derived. If syllabicity is lexically marked, the three-level sonority hierarchy, O→S→V, suffices for practically all syllabification purposes.¹⁴

3. Internal Structure

We will now consider what the possible internal organization of the syllable could look like. (Yip, 2003 gives a similarly sceptical account from a different perspective.) The first important issue to tackle is whether syllable structure is universal or language specific. The question is not whether any two syllables will have identical structure across languages: this would mean that

all languages had to follow the one with the most complex structure.¹⁵ The question is whether variations like those depicted in (9) are possible across languages.¹⁶

(9) Divergent syllable structures



Let us follow the narrower, and therefore more illuminating path: accepting it as an axiom that such variation is not possible. All languages have a uniform syllable structure, potentially one of the above.

In order to evaluate the proposals made about what this uniform structure should look like, we must first collect reasons for assigning internal structure to syllables. A hierarchical system of constituents is posited by linguists to show that some elements are more closely related than others. In the noun phrase *the black cat*, the words *black* and *cat* are within a constituent that does not contain *the* because their relationship is found to be tighter than that of either with the determiner. Also, in feature geometric representations of the internal structure of segments, those features or nodes are listed under a common node that pattern together, that is, spread together, or are deleted together. Syllable internal structure is thus expected to reflect that not all the segments within a syllable are peers of each other.

The prototypical syllable contains a consonant followed by a vowel (CV). Languages in which this type exhausts the syllable type inventory are not instructive of internal structuring of syllables at all. More complex systems allow closed syllables, CVC, and/or long vowels, CVV. In the case of such syllables, subdivisions make sense, in fact, they are obligatory if one accepts that constituents are maximally binary branching (Kayne, 1984; Kaye et al., 1990). If so, the option depicted in (9a) is excluded on theoretical grounds. This 'flat' structure is employed by Kahn (1976) and Clements and Keyser – although the latter also posit a so-called nucleus display, but their nucleus is not a daughter of the syllable (1983: 17). In itself it does not provide much information other than the location of syllable boundaries.¹⁷

The choice between the other two possibilities, (9b), in which the initial consonant enjoys more independence, and (9c), in which the initial consonant and the vowel are more closely related, is not obvious.¹⁸ Arguments can be based on two kinds of phenomena: syllable weight and phonotactic relations between adjacent segments. Vowel length in closed syllables, which may be related to both categories, will be discussed separately.

3.1 Syllable Weight

Syllable weight depends on the number of segments in the second half of the syllable. The vowel always contributes to the weight of the syllable (i.e. it is always moraic). Long vowels, and in some languages syllables closed by a consonant, contribute two moras, thus make the syllable heavy, type (10a). There also exist languages where the weight of syllables depends exclusively on the vowel they contain, any consonant closing the syllable is ignored, type (10b). Let it be mentioned that there exist cases of a third system in which some closed syllables (those closed by a sonorant consonant) are heavy, others (closed by an obstruent) are light (Lithuanian and Kwakwala are well-known examples).¹⁹

(10) Two systems of syllable weight

	light		heavy	example
a.	–V	┌–VC	–VV	English, Hungarian, Cairene Arabic
b.	–V	–VC	–VV	Lardil, Khalka Mongolian

As opposed to syllable-final consonants, those preceding the vowel never²⁰ influence the weight of the syllable. Interestingly, this asymmetry may be used to argue in both directions. One may say that the weightlessness of the syllable-initial consonant sets it off as an independent unit, whereas the fact that a syllable-final consonant is potentially weighty likens it to the vowel, which, recall, always carries weight. Accordingly, the internal structure of the syllable to be preferred is the one in (9b). The intermediate constituent containing the vowel and the consonant following it is called the rhyme. Adherents of this view will then say that only segments dominated by the rhyme may be moraic, those outside the rhyme – in the onset – never carry weight. However, this argument cannot hold for languages where closed syllables are not heavy, type (10b). So one either posits structure (9c) for these languages, at least for closed syllables (McCarthy, 1979: 455), thereby giving up the axiom that syllable structure is uniform across languages, or the weight argument for the rhyme is lost. In other words, reference to syllable weight argues for the constituenthood of the rhyme only in those languages where closed syllables are heavy, type (10a), not in others where they are light, type (10b).

In fact, the selection of structure (9c) could also be justified by the same facts. Under this view, weight is represented by the immediate constituents of the syllable. The weightlessness of the onset now follows from the fact that it always shares a constituent with the vowel, while the syllable-final consonant, the coda, is a separate unit, another mora (cf. Ewen and van der Hulst, 2001: 129).

Unfortunately, the parametric variation of languages with heavy versus others with light closed syllables again cannot be explained in a straightforward way, since in the weighty-coda systems the immediate branching of the syllable node adds weight, in weightless-coda systems it does not. That is, weighty-coda languages can be formalized equally well with both structures (9b) and (9c), but both fail in the case of weightless-coda systems.

Thus, facts about syllable weight do not provide conclusive evidence either for the constituenthood of the onset and the nucleus with the exclusion of the coda, or for that of the nucleus and the coda with the exclusion of the onset.

3.2 Phonotactic Constraints

Let us turn to phonotactic constraints then. The question again is whether there is a significant difference in the intimacy of the relationship of a vowel with the consonants on either of its sides. The right-branching structure in (9b) forecasts that the onset is independent of the rest of the syllable phonotactically, but the coda is not. Against such a view, Clements and Keyser state: 'cooccurrence restrictions holding between the nucleus and preceding elements of the syllable appear to be just as common as cooccurrence restrictions holding between the nucleus and following elements' (1983: 20). They then produce a wholly unconvincing list of English phonotactic constraints to support their cause. As the authors themselves admit, some items on the list mention minor constraints that could easily be taken as accidental gaps. 'Voiced fricatives [...] are excluded before /ʊ/': both voiced fricatives and /ʊ/ are among the less frequent segments of Modern English, that their combination is even rarer is not surprising. But the situation is worse: words like *bivouac*, *casual*, *usual*, *zucchini* do contain the allegedly impossible sequences.²¹ Other constraints on Clements and Keyser's list are more extensive and do not involve exclusively the onset and the following rhyme. For example, the $sC_i\check{V}C_i$ constraint refers to a coda [s] followed by an onset consonant, a vowel, and another consonant that may well be the onset of yet another syllable. That is, this is not a constraint on an onset and the following nucleus/rhyme. In fact, Törkenczy (1994) shows that this is a morpheme structure constraint and has nothing to do with syllables, with respect to this constraint the status of $CC_iV\$C_i$ is the same as that of $CC_iVC_i\$$ (and as $C\$C_iVC_i$ for that matter). Yet other claims Clements and Keyser make are simply false ('in dialects distinguishing /a:/ from /ɒ:/, the sequence /wa:/ is excluded': British English is such a dialect with /wa:/ in *Botswana*, *guano*, *qualm*, *quark*; '/Cl/ clusters are excluded before /ʊ/': *fluoride*, *influence* etc.; '/Cr/ clusters are excluded before /er,or,ar/': *agrarian*, *arbitrary*, *contrary*, *librarian*, *prairie*, *sombrero*, *crore*, *registrar*).²²

Kaye's view is the absolute opposite: he claims that any onset may combine with any rhyme, that is, there are no constraints whatsoever between the

prevocalic part of the syllable and the rest (1985: 290; also cf. Pike and Pike, 1947; Selkirk, 1982). Even if some languages do exhibit onset–nucleus constraints (like *[wu], *[ji]), the robust pattern is that these are few and far between. This constitutes an argument against the left-branching structure in (9c). But it does not automatically justify the right-branching structure in (9b). If one is to argue for that, one has to show that the vowel and the consonant at the end of the syllable are related.

In fact, the interaction between a vowel and the following consonant is not as common as tacitly assumed by proponents of structure (9b). The coda constraints found in languages typically constrain the types of segment that can occur in coda position, irrespective of the preceding vowel, in fact, often depending on the following consonant.

Fudge notes that in English certain consonant clusters can only be preceded by short vowels, thus long vowels and diphthongs do not appear before noncoronal nasal+stop or stop+stop clusters (1969: 272f.). Many other languages restrict the length of the vowel in a closed syllable; note, however, that this does not effect the quality of the vowel, but its quantity.²³ That is, it is not necessary to include the vowel and the following consonant in a single constituent to explain why no consonants may occur after a long vowel. We will return to this point in Section 3.3.

Examples for genuine qualitative changes happening to a vowel because of the following consonant are breaking/broadening in English or nasalization in French. Consider the data in (11) taken from Giegerich (1992: 55, 63). They are somewhat simplified to illustrate the point. S stands for Scottish Standard English (SSE), A for General American (GA), and B for Standard Southern British English (SSBE).²⁴

(11) Plain and pre-R vowels in three accents of English

a. plain				b. pre-R			
	S	A	B		S	A	B
shoot	u	u	u	sure	u	ʊ	ɔ
spoke	o	o	o	sport	o	ɔ	ɔ
shot	ɔ	ɑ	ɒ	short	ɔ	ɔ	ɔ
butt	ʌ	ʌ	ʌ	word	ʌ	ɜ	ɜ
bid	ɪ	ɪ	ɪ	bird	ɪ	ɜ	ɜ
head	ɛ	ɛ	ɛ	heard	ɛ	ɜ	ɜ

All three accents exhibit six different vowels before consonants other than [r], as shown in (11a). In pre-R position, however, only SSE has the same vowels, in GA we find three different vowels, in SSBE the contrast reduces to two only, (11b). This difference is accompanied by the fact that in SSE [r] is a full-fledged

consonant, trill or flap or fricative (Wells, 1982: 411), while in SSBE it merges with the preceding vowel and disappears altogether. In General American nonprevocalic R is realized as an r-coloured vocoid (Wells, 1982: 490). Thus, R influences the preceding vowel only if it is vocalized, that is, when it becomes part of the nucleus, but not when it is consonantal, that is, when it is in the coda. (Also cf. Harris, 1994: 257ff.) French front vowels are lowered by a following nasal, but again only when the nasal vocalizes and merges with the preceding vowel (cf. *fin* [fɛ̃]~*fine* [fin] 'fine masc.~fem.', *un* [œ̃]~*une* [yn] 'one masc.~fem.').

Without an exhaustive survey of vowel-consonant sequences across languages, we may risk the claim that phonotactic constraints are no more frequent between nucleus and coda than between onset and nucleus (Harris, 1994: 64). I agree with Clements and Keyser, after all: there are no significant phonotactic constraints between consonants and vowels. Such constraints hold between members of diphthongs, long vowels, as well as between members of consonant clusters, irrespective of whether they are in the same syllable, or not.

3.3 Closed Syllable Shortening

Another apparent argument for structure (9b), the onset versus rhyme split of the syllable, is the phenomenon of closed syllable shortening. In a number of languages, a syllable is either closed or has a long vowel, but not both. Italian non-final stressed syllables exemplify the phenomenon,²⁵ some data are given in (12).

(12) Closed syllable shortening: Italian stressed syllables

a. open syllable → long vowel

fā.ce 'torch', fā.ma 'fame', fā.me 'hunger', fā.ro 'lighthouse',
fā.se 'phase', fā.ta 'fairy', fā.va 'bean', fi.bra 'fiber'

b. closed syllable → short vowel

fāb.bro 'smith', fāg.gio 'beech', fāl.ce 'sickle', fāl.co 'falcon',
fān.go 'mud', fān.te 'pedestrian', fās.to 'luxury', fīr.ma 'signature'

The syllable boundary is marked in the words to show that when it follows the vowel immediately, the vowel is long, when it does not, the vowel is short. This distribution can be explained by claiming that the rhyme constituent must contain maximally (in the case of shortening in closed syllables) or exactly (in the Italian case discussed here) two segments.

However, structure (9c) explains such a restriction just as well: this tree contains a single position following the onset and the first part of the nucleus,

thus there is no place for a coda consonant after a long vowel in this tree, as (13a) shows.

(13) Left- and right-branching syllable trees and closed syllable shortening



In fact, the onset–rhyme model cannot exclude a closed syllable containing a long vowel by reference to binary branching alone: as (13b) shows, a rhyme can contain three segments *and* observe binary branching (cf. Harris, 1994: 163f.). To exclude this structure – and explain closed syllable shortening – Charette (1989: 161ff.) borrows the notion of c-command from syntax, while Kaye et al. (1990: 199) claim that the three-member rhyme constituent of (13b) is impossible because the head cannot properly license all its dependents.²⁶

Structure (9c) – as (13a) shows – falls short of explaining the phonotactic constraints evidently holding within a complex nucleus: diphthongs in a language are typically a small subset of all the possible combinations of short vowels. But then, as we have seen, phonotactic constraints hold even between segments that do not belong to the same syllable: consonant clusters flanking a syllable boundary often are so related.

To conclude, we have not found compelling evidence for either the onset versus rhyme split or the onset-cum-nucleus versus coda split within the syllable. The theoretician now has two options: either he reverts to the flat syllable tree of (9a), which undesirably violates binarity, or he asks whether the syllable node exists at all. What phonotactic constraints call for is primarily vowel–vowel and consonant–consonant communication, which is clearly independent of syllables. By dispensing with the syllable as a constituent in the prosodic hierarchy, we do not automatically make syllable-related explanations of phonological phenomena impossible. In what follows, I will discuss an alternative model of syllable structure that lacks any hierarchical structure, but is still capable of expressing most of what standard syllable trees can.

4. Vowelless Syllables

It has already been hinted at that word edges exhibit peculiar behaviour with respect to syllable structure. For example, two obstruents (excluding the weirdly

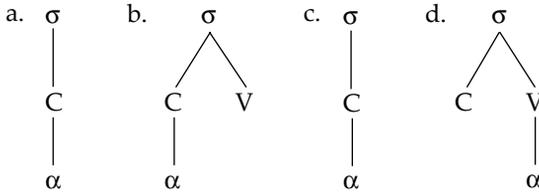
behaving [s]) cannot be adjacent within a syllable except when they are at the edge of a word. The definition of the universal syllable in (4) predicts that such a cluster will be separated by a syllable boundary and this does not depend on the position of the cluster within the word. That is, words like the following must contain a syllable boundary at the dot: *pas.t*, *sof.t*, *fac.t*, *s.top* etc. Note that now two obstruents are not adjacent within a syllable even at the edge of a word. Recall that in some languages after such a syllabification epenthesis ensues (e.g. in Yiddish [g.ɔ.ɾɪm]>[gɪ.ɔ.ɾɪm]). But in the above English examples epenthesis does not occur.

In the case of another prosodic unit, the foot, it is accepted that word boundaries may not coincide with foot boundaries. A foot may span across a word boundary (e.g. the first foot in |*Sám's a|lône*|), and it may also occur that we only find the second half of a foot at the beginning of a word, for example, in *a|lône*, where the second syllable initiates a foot, while the first, unstressed syllable lacks a head. Such a foot is called a degenerate foot.

Since syllables are also prosodic units, and thus their segmentation should be largely independent of semantic units (i.e. words), I accept Aoun's (1979) proposal, adopted and elaborated by Kaye (1990a), that such vowelless (i.e. headless) syllables are degenerate syllables.²⁷ Thus, the last consonant of *pas.t* forms a degenerate syllable that lacks a vowel. Claims like 'there are no well-formed syllables in any language that lack an overt nuclear segment on the surface' (Hayes, 1989a: 286), or 'in all languages, syllable edges correspond with word/utterance edges' (Blevins, 1995: 209) have no theoretical status, they describe beliefs. One has to support such a claim with arguments. The onset of a syllable may be empty in a large set of languages, it is not clear why we should accept without argument that the same status is denied of nuclei. One could argue that the nucleus is the head of the syllable, but headless feet are accepted and headless structures occur abundantly in syntax, too. If the sentence is analysed as a complementizer phrase, the head complementizer is very often missing from the surface. Another possible argument could be that a syllable is unpronounceable without a nucleus. But in words, nucleusless syllables always co-occur with others that have a nucleus: words always contain some syllabic segment, it is clitics that do not and thus need the phonological support of an adjacent word. In fact, as we will see, the number of consequent nucleusless syllables is strictly constrained in languages. The general aversion towards vowelless syllables stems from the fact that the view of the syllable as a wave of sonority seems impossible without the peak of sonority that the vowel represents. The definition of the universal syllable in (4) shows that this is not the case: the empty vowel between two obstruents can be detected, just like the empty consonant between two vowels that create hiatus, as both constitute segment pairs of equal sonority.

There are two senses of a syllable being vowelless (or consonantless), as shown in (14), where α stands for any segment.

(14) Emptiness in a syllable



Both (14a) and (14b) represent a degenerate syllable, but while the first syllable contains a sole consonant (this is Aoun's conception), the second is a prototypical syllable, it is just that its nucleus is not associated with any melodic material: it is empty. The same is true of the two potential representations of the onsetless syllable in (14c) and (14d). In what follows, I will adopt a version of representations (14b) and (14d).²⁸

As mentioned in note 27, the independence of word-final consonants could be solved without the introduction of degenerate syllables, by labelling them extra-metrical. Myers says: 'An extra-metrical element is inaccessible to the phonology, that is, it can neither condition nor undergo phonological operations' (1987: 487). In a note, Myers adds that the inaccessibility refers only to the extra-metrical element, not to its features. This is necessary to explain the phonotactic constraints holding between two word-final consonants (cf. Harris, 1994: 66ff.). Still, if a word-final consonant is extra-metrical then we do not expect to find $-\check{V}C\#$ sequences in English, since word-final stressed short vowels do not occur. This paradox does not arise with the degenerate-syllable analysis.

In what follows, we will extend the distribution of degenerate syllables in an extreme way: any consonant will be seen as the 'onset' of a 'syllable' either with or without a vowel.

5. Syllables without Constituents

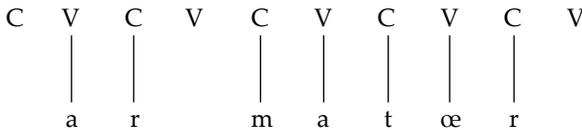
Let us begin with quoting²⁹ the French academician Charles Pinot Duclos (1704–1772):

Pour distinguer la syllabe réelle ou physique, de la syllabe d'usage, il faut observer que toutes les fois que plusieurs consonnes de suite se font sentir dans un mot, il y a autant de syllabes réelles qu'il y a de ces consonnes qui

se font entendre, quoiqu'il n'y ait point de voyelle écrit à la suite de chaque consonne: la prononciation suppléant alors un *e* muet, la syllabe devient réelle pour l'oreille, au lieu que les syllabes d'usage ne se comptent que par le nombre des voyelles qui se font entendre et qui s'écrivent. [. . .] Par exemple, le mot *armateur* seroit en vers de trois syllabes d'usage, quoiqu'il soit de cinq syllabes réelles, parce qu'il faut suppléer un *e* muet après chaque *r*; on entend nécessairement *aremateure*. *Bal* est monosyllabe d'usage, et dissyllabe physique. (Arnauld and Lancelot, 1803: 407f.)

Whenever it is felt that consonants are adjacent in a word, each consonant constitutes a syllable (réelle), even if there is no vowel following it. The traditional sense of the syllable (la syllabe d'usage) counts only pronounced vowels. Duclos claims that any consonant that is not followed by a vowel constitutes a degenerate syllable: the superficially trisyllabic word *armateur* is in fact to be syllabified as five syllables *a.r.ma.teu.r*, shown in (15) in current nonlinear terms.

(15) The representation of *armateur*



As for vowel clusters, Duclos says:

A l'égard de la diphtongue, c'est une syllabe d'usage formée de deux voyelles, dont chacune fait une syllabe réelle, *Dieu, cieux, foi, oui, lui*. (Arnauld and Lancelot, 1803: 408)

So, just like consonant clusters, diphthongs also form two underlying 'syllables', that is, two vowels separated by an empty consonant.³⁰ Almost two and a half centuries later,³¹ Lowenstamm (1996) arrives at the same conclusion: the phonological skeleton consists of strictly alternating consonantal and vocalic positions, as shown in (15). Note that if the simplest sonority hierarchy (C→V) is applied to the universal syllable in (4), then exactly this 'syllabification' follows. As regards their position on the skeleton, two vowels or two consonants are never adjacent. The model also lacks any branching above the skeleton: syllabic constituents are missing. This conception of prosodic organization solves a number of issues left unresolved above: whether the onset or the coda belongs with the nucleus, or whether syllabicity is lexically marked.

However, several other questions arise in connection with this extremely simplified syllable structure. Some of them are listed in (16), and discussed in the following sections.

- (16) Questions about the strict CV view of the skeleton
- a. how are empty V positions controlled?
 - b. why is the length of consonant clusters limited in natural languages?
 - c. why is the length of vowels limited in natural languages?
 - d. why are some consonant clusters possible, and others not?
 - e. how are long vowels/diphthongs and hiatus distinguished?

5.1 Consonant Clusters

If consonant clusters do not form constituents whose size can be limited, there has to be some alternative method of constraining them. In a syllable-tree model, the size of a syllabic constituent containing consonants may be controlled by the number of branches that it may have. In the strict CV approach the same effect is achieved by controlling the number of subsequent empty vowels. Emptiness is marked:³² the occurrence of an empty vocalic position must be sanctioned. Kaye et al. claim that it is the pronounced vowel of the next syllable that sanctions the nonpronunciation of an empty vowel by a relation they call (proper) government (1990: 219). The pairs in (17) taken from various languages show that the appearance of the vowel indeed depends on the appearance of the following vowel. (Sources of the data: Brockhaus, 1995: 214 for German; Gussmann, 2007: 185 for Polish; Kaye, 1990b: 140 for Moroccan Arabic; Kenstowicz, 1994: 129 for Somali; my own competence for the other languages.)

- (17) Vowel–zero alternations across languages
- a. English: *simple* [sɪmpəl] ~ *simpler* [sɪmp•lə]
 - b. French: *j'amène* [ʒamɛn] 'I bring' ~ *amener* [am•ne] 'to bring'
 - c. Latin: *pater* 'father' ~ *pat•ris* 'gen.'
 - d. German: *Eben* 'level' ~ *Eb•nung* 'levelling'
 - e. Polish: *łeb* 'head' ~ *ł•ba* 'gen.'
 - f. Hungarian: *majom* 'monkey' ~ *maj•ma* 'his/her monkey'
 - g. M Arabic: [tan k•dɪb] 'I lie' ~ [tan kɪd•bu:] 'we lie'
 - h. Somali: [nɪrɪg] 'baby female camel' ~ [nɪr•go] 'pl.'

In each morphologically related pair in (17) the vowel in ___C# position is pronounced, but when a suffix provides a pronounced vowel after it, that is, it comes to be in ___CV position, it remains silent (marked by a bullet). English

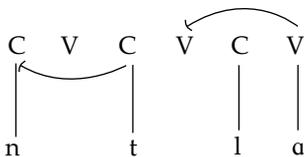
It is also to be observed that no vocalic melody is associated to the template in (18). Recall the crudest sonority hierarchy mentioned in Section 2.3: C→V. Dienes and Szigetvári (1999) claim that this hierarchy represents the top and bottom of the ideal sonority wave in speech: loudness (V) and silence (C), it is the alternation of these two that the skeleton encodes. The prototypical consonants, plosives, are a brief period of silence, the prototypical vowel is a peak of loudness. That is, a V slot is loud, and a C slot is silent, without any melody attached to it. The effect of government can now be generalized as in (19).

(19) The effect of government

Government spoils the inherent property of its target.

The loudness of a governed V position is spoiled: such a vowel remains silent. The ability of a vowel to govern and license is also extinguished by government. For a C position government tends to spoil its silence: the first consonant in a C-to-C governing domain is typically a sonorant. Languages may be more tolerant in the types of consonant cluster they allow – besides the unmarked sonorant+ obstruent cluster others may occur in some languages –, but then C-to-C government is not the only configuration for surface consonant clusters, as (18b) shows. Thus, word finally we expect only consonant clusters that are created by C-to-C government (since there is no V position to govern the V enclosed within the cluster). We also have a means of modelling a word-internal three-member consonant cluster now,³⁸ as shown in (20): this is the only way two successive empty V positions can be silenced, since the first empty V position cannot be governed (silenced) by the second. (A further option is excluded in (24).)

(20) Word-internal three-member consonant cluster



Thus, consonant clusters are of two types: those that are created by government between two C or between two V positions. The differences between the two types are discussed in detail elsewhere (Szigetvári, 2007).

Now that any C is followed by a V, we seem to lose the original generalization about syllable-initial and syllable-final consonants, that is, onsets and codas. In fact, the two environments can still be captured: an onset is a C position

followed by a pronounced V position, a coda is one that is not followed by a pronounced V position.

5.2 Vowel Clusters

If the representation of long vowels and diphthongs involves an empty C position between the two V parts, these and hiatus become identical. In a hierarchical framework, they are distinguished by a branching structure, (21a) versus (21b), or also an empty C slot, an onset, (21a) versus (21c). (Since they may here branch, I use the labels O and N, not C and V, for the consonantal and the vocalic parts of the representation.)

(21) Diphthong versus hiatus (traditional view)



In the strict CV model, neither of the first two options are available, since there are no branching structures, and any two V positions are separated by a C position. Long vowels/diphthongs and hiatus cannot thus be distinguished by their skeletons, some further mechanism is needed.

Prosodic licensing is a widely accepted notion in phonology (Itô, 1986; Goldsmith, 1990; Harris, 1997). It is meant to be an explanation of why onsets allow more contrasts than codas, or stressed nuclei than unstressed ones: the former are better licensed than the latter. It is also consensual that vowels license the preceding consonant (e.g. Kaye et al., 1990; Charette, 1992; Harris, 1997; Steriade, 1999c; Cyran, 2010). The strict CV model applies licensing, besides government, as a relation between positions on the skeleton: any V position that is pronounced licenses the C position before it (Ségéral and Scheer, 2001). Following the standard assumption, the effect of licensing is given in (22).

(22) The effect of licensing

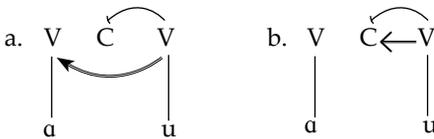
Licensing supports the melodic content of the licensed position.

However, contrary to standard assumptions, licensing – as Ségéral and Scheer conceive of it – is not a *sine qua non* for all portions of the representation: C positions that are not followed by a pronounced vowel are not licensed. This results in their reduced capacity to maintain melodic elements (i.e. contrasts),

but does not necessarily inhibit expression of the given position. That is, licensing as conceived of here is not prosodic licensing (sanctioning the presence of a position) or government licensing (sanctioning a consonant cluster), but exclusively autosegmental licensing (sanctioning melodic elements associated with a skeletal position).

It is also important to see that a licenser is not necessarily more prominent than its licensee – unlike, for example, Harris (1990, 1997) sees it –, it is merely a pronounced vocalic position. Dienes and Szigetvári (1999) contend that two pronounced V positions surrounding an empty C position may be lexically specified to be in a licensing relationship.³⁹ This V-to-V licensing⁴ domain defines long vowels and diphthongs, as shown in (23).

(23) Diphthong versus hiatus (strict CV view)



Like C-to-C government, V-to-V licensing is also possible only across an empty position. But since a C position is silent unless it is forced to be pronounced (by government or by being associated with melodic elements), there is no need to sanction empty C positions. Thus, complex nuclei and hiatus are only different in the presence and absence of licensing between two neighbouring V positions.

The standard explanation for the nonexistence (or marginality) of three-way vowel length contrast is the absence of ternary branching syllabic constituents, nucleus in our case. The framework introduced here seems to allow such structures, as shown in (24).

(24) Ternary clusters



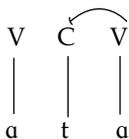
To encode the observation that such vowel and consonant strings do not occur in natural language, the possibilities above must be excluded. As for (24b), one could claim that being governed deprives a position from its governing capacity – as was claimed for V positions above –, but licensing certainly does not do so. If it did, a consonant before a long vowel would be unlicensed, which is an absurd possibility: it would entail that prevocalic consonants differ in their

readiness to lenition depending on the length of the vowel. Instead, Dienes and Szigetvári (1999) call C-to-C government and V-to-V licensing burial domains (burying the enclosed empty V or C position), and stipulate a constraint inhibiting burial domains from overlapping. This shows that abandoning constituency prevents the analyst from using well-established notions like binary branching in syllabic constituents. On the other hand, it is a distinct advantage of the treeless model that it creates relations between segments that could not be related in syllable trees. Recall, for example, that phonotactic constraints link a coda and the following onset. This sequence cannot be made into a syllabic constituent in any traditional hierarchical framework. In the present treeless model, however, this relation is not different in kind of that holding a long vowel or diphthong together. Although Kaye et al.'s (1990) theory posits very similar relations between skeletal positions as here, it also contains syllabic constituent trees. Since the relations and the constituents encode the same facts, Takahashi (1993) rightly points out that one of the two devices is superfluous.

5.3 Intervocalic Consonants

Recall, V-to-V government may be established between a pronounced and an empty V position, that is, one which is not associated with any melody. The nonalternating vowels in (17) are lexically associated. The following vowel cannot govern a vowel associated with melody (Ségéral and Scheer, 2001). In this case, it governs the consonant between the two vowels, as shown in (25).

(25) Intervocalic consonants are governed

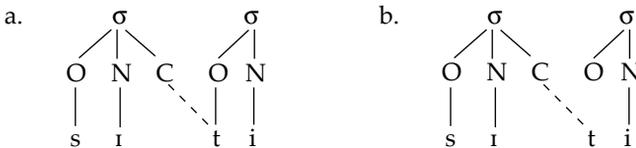


Since government spoils the inherent properties of its target (cf. (19)), the muteness of the C position in our case, the prediction is that intervocalic consonants will lose their muteness and become more vowel-like, more sonorous. This status of intervocalic consonants enables us to evade debatable notions of syllable theory: coda capture, which may be implemented either as ambisyllabicity or as complete resyllabification.

It is a well-documented fact that consonants in the coda of a syllable are 'weaker' than those in the onset: the latter are (better) licensed by the vowel that follows, the former are not (so well licensed), since no vowel follows them. This is manifested in two interrelated phenomena. Languages may limit the set of

consonants – or rather the set of phonological features – that is allowed to appear in codas (by so-called coda constraints, cf. Itô, 1986; Borowsky, 1986), but similar restrictions are very unusual for onsets. Coda consonants are also more prone to undergo lenition than onset consonants. Intervocalic consonants, however, are often also subject to lenition: flapping in several accents of English is an example. To synchronize lenition sites and syllabic affiliation several authors have proposed that core syllabification (where an intervocalic consonant is syllabified to the right as an onset) is followed by a process that assigns a (post-tonic) intervocalic consonant to the coda (or alternatively to both the coda and its original onset), that is, a (stressed) syllable captures the onset of the following syllable to make it its own coda (e.g. Kahn, 1976; Selkirk, 1982; Gussenhoven, 1986; Rubach, 1996). The change is illustrated by the ambisyllabicity, (26a), and the resyllabification, (26b), of the intervocalic [t] in *city*.

(26) Ambisyllabicity and resyllabification in *city*



The idea of coda capture is theoretically uneasy. Ambisyllabicity – introduced by Hockett (1955: 52) as interlude – itself is highly problematic (cf. Selkirk, 1982; van der Hulst, 1985; Harris, 1994; Jensen, 2000), challenging the idea of proper bracketing. The total resyllabification of an onset consonant into the coda of the preceding onset works against onset maximalization, an apparently universal principle. Furthermore, the reasoning is circular: intervocalic consonants are resyllabified only to explain why they undergo lenition.⁴⁰

A proposal to solve the problem of ambisyllabicity is to analyse such consonants as virtual geminates, a consonant that is represented as it were a geminate, but pronounced short. For example, van der Hulst (1985) argues that an intervocalic consonant following a short/lax vowel in Dutch is representationally a geminate, hence all such vowels are in a closed syllable. Now the consonant belongs to two skeletal positions, which can belong to two syllables respecting proper bracketing. The lenition facts of English, however, do not allow such an analysis. There is no difference in flapping between the two types of [t], that following a long/tense and a short/lax vowel: *treaty* and *city* exhibit the same kind of flapping. In addition, if the latter had a virtual geminate, geminate integrity would make lenition outright unlikely. So the environment posited to condition lenition would be one where lenition were least likely.

To unify the two types of lenition environment, Harris (1997) advocates an intricate system of lenition paths starting out from a stressed nucleus, which

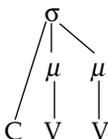
licenses the unstressed nucleus in its foot, both nuclei licensing their onsets, and onsets licensing the preceding coda. The further down the licensing path, the weaker a position: this renders the onset of an unstressed syllable weaker than that of a stressed syllable. That is, the weakness of a coda compared to an onset is a consequence of licensing relations, not of syllabic affiliation. There is thus no need to resyllabify anything. However, Harris's model predicts (i) that the onset of *any* unstressed syllable will be weak and (ii) that intervocalic lenition is restricted to nonpretonic onsets. Neither prediction is right: while the [t] in *city* is flapped, that of *filter*, *mister*, *actor* is not. In fact, it is only intervocalic⁴¹ [t] and [d] that is flapped in English, exactly as (25) predicts. Furthermore, although English flapping happens only before an unstressed vowel, in other systems, for example, in Tuscan Italian, intervocalic stops spirantize irrespective of stress (e.g. *abete* Standard Italian [a'βete], Tuscan Italian [a'βe:θe]; Marotta, 2008: 242). Such a language cannot be analysed in a framework that blames lenition on the distance of the licensing stressed vowel. The government-based lenition of consonants shown in (25), on the other hand, predicts sonorization of intervocalic consonants irrespective of stress, catering for the Tuscan pattern, but not for the English one. Ségéral and Scheer (2008: 507ff.) and Szigetvári (2008: 581) propose two different means of accounting for the absence of lenition pretonically. Although this way an extra mechanism is involved in the analysis, the same mechanism is needed anyway to explain why syncope is also impossible before a stressed vowel (*memory* ['mɛm(ə)ri] versus *memorize* ['mɛm*(ə)raɪz]),⁴² since – as is claimed here – both the lenition of an intervocalic consonant and syncope are the effect of the government coming from the following vowel.

5.4 Syllable Weight

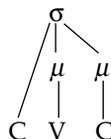
As we have seen in Section 3.1, the internal constituency of syllables cannot adequately describe the two basic weight systems of (10). Replacing constituents like nucleus and coda with moras (Hock, 1986; Hayes, 1989a) was an answer to this shortcoming. Theories applying moras distinguish the two weight systems of (10) by assigning a mora to both the vowel and the consonant following it, (27b), or only to the vowel, (27c).

(27) Weight indicated by moras

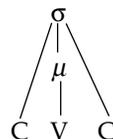
a. long V heavy
=(10a, b)



b. coda heavy
=(10a)



c. coda light
=(10b)



A syllable containing a long vowel is always heavy, since vowels are invariably assigned a mora, (27a). The variability is in whether the coda is or is not assigned a mora. Scheer and Szigetvári (2005) claim that it is arbitrary that a mora may be assigned to a coda consonant, but not to one in the onset. This fact, however, follows from the strict CV model, where exclusively V positions are moraic. The difference between (10a) and (10b) depends on whether empty V positions are counted or not. Since what is traditionally called an onset consonant is always followed by a pronounced vowel, it is always in a CV pair that is moraic, its weight being attributed to the vowel. Codas, on the other hand, are followed by an empty vowel. If this empty V position is counted (in weighty-coda languages), the coda consonant appears to carry a mora. Actually C positions are never moraic, when they appear to be, the illusion arises because the following empty V position participates in the mora count. This solution is less arbitrary because the same type of objects, V positions, contribute to syllable weight. It is not conceivable that in some language prevocalic consonants are moraic, but others are not. This nonattested scenario is not excluded by the moraic model. (In (28) unpronounced positions are symbolized by a lower-case letter, moraic V positions are underlined; the three representations parallel those of (27).)

(28) Weight indicated by CVs

a. long vowel =(10a), (10b)	b. empty V counts =(10a)	c. empty V ignored =(10b)
C <u>V</u> c <u>V</u>	C <u>V</u> C <u>v</u>	C <u>V</u> C v

6. Conclusion

This chapter argues that (i) syllable structure is a useful notion in the description and explanation of phonological phenomena, but (ii) the representation of syllable structure by constituents visualized in trees runs into difficulties. Instead, a severely reduced model was offered applying a skeleton of strictly alternating consonantal and vocalic positions, representing muteness and loudness, respectively. The relations between segments are modelled by two forces, licensing and government. It is shown that this framework establishes relations only between segments that do exhibit phonotactic constraints, irrespective of whether they are tauto- or heterosyllabic. We get a noncircular explanation for why intervocalic consonants may undergo lenition. The strict CV model also offers an explanation for why onsets do not contribute to the weight of the

syllable, while codas potentially do, instead of simply encoding this fact in the representation.

7. Acknowledgements

I am indebted to two anonymous reviewers, András Cser, Ádám Nádasdy, and Péter Siptár for comments on this paper. Errors are my responsibility.

8. Notes

1. Think of Regina Spektor's *Consequence of sounds*, for example.
2. McCawley (1965) is a notable exception.
3. In this chapter, I use terms like 'onset', 'coda', or 'syllable' in a descriptive sense, as it will become clear, their theoretical status is not firm.
4. The alternative solution is to insert a vowel to save such a consonant. This happens in *rhythm*, where //rɪðm// (cf. *rhythmic* [rɪðmɪk]) surfaces as [rɪðəm].
5. The distinction between analytical and nonanalytical affixation is roughly equivalent to that between level-1/lexical and level-2/postlexical affixation in lexical phonology. The crucial difference is that lexical affixation is denied any phonological relevance. In modern terminology, there is no spell-out at the lexical level (Scheer, forthcoming).
6. Kahn clearly states that his theory is about syllables 'on the phonetic level' (1976: 20), Lowenstamm talks about 'underlying' syllables. What follows is that superficial syllabification cannot be achieved by an algorithm.
7. (4a) is equivalent to Selkirk's sonority sequencing generalization (1984a: 116).
8. There are no languages with syllable-initial consonant clusters, but no closed syllables. Kaye and Lowenstamm (1981) discuss implicational relations between syllable types; Blevins (1995) discusses syllable typology; and Piggott (1999) consonant cluster typology.
9. Note that with consonants this problem does not occur: intervocalic geminate consonants are heterosyllabic. As will be argued, word-final consonant clusters do not belong to the same syllable.
10. Provided that these 'nuclear' glides are not categorized as *C*, in which case they are less sonorous than the preceding vowel.
11. Although there is a tradition of positing a ranking among obstruents along the voiceless→voiced and/or plosive→fricative dimensions (e.g. Anderson and Ewen, 1987), this is to explain lenition and is unnecessary for the purposes of sonority sequencing in syllables, as Clements (1990) argues. Fricative+plosive clusters – especially with [s] – may occur both initially and finally (e.g. *stop/post*), or in reverse order (e.g. *ask/axe*). In English this is clearly not controlled by sonority sequencing. In Dell and Elmedlaoui's (1985) analysis of Imdlawn Tashlhiyt Berber, sonority distinctions within obstruents are required to derive the syllabicity of consonants. Guerssel (1990) and Coleman (1999), however, question the existence of syllabic obstruents and claim that reduced vowels occur next to these.
12. Exceptions (like *album* *[-bm]) are probably not related to sonority. Another constraint allows syllabic [r] – and nonsyllabic [r] for that matter – only before a syllabic segment.
13. The last two items in (7b) are given in a rhotic pronunciation, which has syllabic [r] in this position.

14. More elaborate hierarchies are needed if we wish to account for phonotactic constraints, for example, in English [pr] and [pl] exist syllable-initially, but [pn] does not. Assigning a sonority index to segments, or segment classes, would enable the analyst to formulate conditions on the necessary sonority distance between segments in syllables. But even these are not enough, a number of constraints (e.g. [pl] versus *[tl], *[pw] vs. [tw]) cannot be explained in this way (cf. Selkirk, 1984a).
15. Quite to the contrary: at the end of this chapter it will appear that all languages follow the simplest structure, CV.
16. These three cases exhaust the possibilities unless association lines are allowed to cross.
17. Kahn chooses the nonlinear syllable tree of (9a) to provide for ambisyllabic consonants. With syllable boundaries inserted between segments ambisyllabicity would either be impossible – V\$C\$V indicates a syllable comprising a single consonant –, or, with labelled bracketing, unconstrained – besides [ₐ . . . V[ₑ C]ₐ V . . .]ₑ, the bracketing [ₐ . . . V[ₑ CC]ₐ V . . .]ₑ would also be possible. The tree representation, together with the no-crossing-lines constraint allows the option of being ambisyllabic only for single consonants.
18. Nevertheless, the structure in (9b) is significantly more popular among theorists. McCarthy is one who suggests a tree like (9c) to explain the extrametricality of word-final consonants in Cairene Arabic (1979b: 453).
19. Ewen and van der Hulst (2001: 135) tentatively suggest that Dutch may exemplify a fourth type, in which closed syllables are heavy, but syllables with a long vowel are light. Kenstowicz claims that such languages do not exist (1994: 428). We will return to this issue in note 23.
20. If not never, then so marginally that the counterexamples (e.g. Everett and Everett, 1984; Davis, 1988) can, for the time being, be safely ignored as potential misanalyses (cf. Takahashi, 1999; Goedemans, 1996).
21. Northern England accents that have /ʊ/ where southern accents have /ʌ/ (e.g. *vulgar* [vʊlgə], *result* [rɪzʊlt], *thus* [ðʊs]) show clearly that this ‘constraint’ is the result of a historical process in the south that weeded out most occurrences of /ʊ/, that is, the scarcity of such sequences is due to an independent sound change, not their ill-formedness.
22. In English, onset–nucleus constraints all involve consonant+glide onsets: any vowel can follow a glide, but Cj can only be followed by [u:] (and its r-influenced version), while Cw – where C is not [s] – cannot be followed by [u:] (or [au]) either, which incidentally is the continuation of Middle English [u:].
23. Claiming that it is tense vowels that are excluded here has a highly undesirable consequence: syllable weight would thus be dependent not on the quantity but the quality of the vowel. In fact, this is the solution to the weird pattern of Dutch: the stress placement algorithm sees closed syllables as heavy, long-vowelled ones as light, which is an absolutely unexpected pattern. If ‘long’ vowels are, in fact, tense and short, Dutch fits into the well-established categories described in (10) (Ewen and van der Hulst, 2001: 134f.).
24. Formerly known as Received Pronunciation (RP).
25. The Italian data may also be analysed as open syllable lengthening: the vowel of an open syllable is lengthened so that stressed syllables are always heavy. Crucially, the vowel of a closed syllable is not lengthened.
26. Note that the trouble the onset–rhyme model faces stems from labelling both the nucleus and the rhyme. The original tree in (9c), without these labels, is just as restricted. In it the rhyme can contain two segments: either two vowels, or a vowel and a consonant.
27. Aoun restricts the use of degenerate syllables to the last, ‘extra’ consonant of super-heavy (-VVC or -VCC) syllables, while Kaye extends the notion to any word-final

- consonant. The fact that a word-final consonant does not influence the weight of the last syllable in many languages, English among them, was formalized by labelling these consonants extrametrical (Hayes, 1982). The phonotactics of word-final consonants also shows their independence (Itô, 1986; Myers, 1987; Borowsky, 1989). Harris argues extensively that the set of English word-final consonant clusters is remarkably similar to that of intervocalic, *heterosyllabic* clusters (1994: 66ff.), Harris and Gussmann (1998) show that the western tradition of syllabifying word-final consonants as a coda is probably wrong.
28. The syllable in (14d) is not considered degenerate, just like a monosyllabic foot is not a degenerate foot. Both are, nevertheless, marked structures compared to a CV syllable or a strong–weak foot.
 29. Thanks to András Cser for calling my attention to this comment.
 30. Perhaps Duclos is not right about the representation of French light (monomoraic) diphthongs, his idea is still very *avant-garde*.
 31. Although I here quote from the 1803 edition, Duclos's comments on the Port Royal grammar were first published in 1754.
 32. Kaye and Lowenstamm (1981) and Cairns and Feinstein (1982) discuss markedness and syllable structure; Szigetvári (2006) discusses the relation of markedness and complexity.
 33. Whether it is sonority or some other property of sounds – charm (Kaye et al., 1985, 1990) or complexity (Harris, 1990) – that controls their sequencing is a question I do not address here.
 34. C-to-C government is drawn below the intervening V, because this type of government is only possible across an empty V: the arrow graphically inhibits association to melody.
 35. Amimi (1996) shows that the distribution is more complicated with competing alternants. He claims that the basic CCVC template – which lacks C-to-C government – may surface as CVCC, if C-to-C government is possible.
 36. These facts also imply that the governing relation between skeletal positions is not part of the template, but depends on the melodic content associated with the given position.
 37. The notion of C-to-C interaction has alternative roles in later analyses. Cyran and Gussmann (1999), for example, call it interonset government, but they apply it in the reverse direction. Scheer (1996, 2004) applies right-to-left infrasegmental licensing. These models differ from Kaye's implementation in that they silence the vowel within a rising-sonority cluster, that is, a branching onset. C-to-C government as defined here is thus a return to Kaye's original version.
 38. Four-member clusters, as in English *i[nstr]ument*, *e[ksp]ain*, cannot be dealt with. These all involve [s], the *enfant terrible* of phonotactics (cf. Kaye, 1996).
 39. Yoshida (1993) and Ritter (1995) analyse the long vowels of Palestinian Arabic and Hungarian, respectively, as constituting a left-to-right governing domain enclosing an empty onset.
 40. The argument that coda capture is also necessary because words, and hence syllables, do not end in a short vowel in English is a weak one: if this were true, a word like *ferry* could not be syllabified at all. Neither [fɛ], nor [fɛr] is a possible word-final sequence in English, if the latter were syllable final, it would be a unique case where [r] occurred in the coda.
 41. The [t] in *party* and *twenty* is intervocalic following a rhotic and a nasalized vowel, respectively.
 42. The absence of pretonic syncope cannot be explained by stress clash avoidance, since syncope does not occur even if it would not result in stress clash: for example, *hùllab*(a)lòo, mèthod*(o)lógical*.

5 Stress

Diana Apoussidou

Chapter Overview

1. Introduction	95
2. Phonetic and Phonological Aspects of Stress	96
3. Modern Greek	100
4. Simulating the Acquisition Process	111
5. Discussion and Conclusion	125
6. Acknowledgements	126
7. Appendix	126
8. Notes	134

1. Introduction

Stress in linguistics is the relative prominence of one syllable in comparison to other syllables. It is part of a language's prosody and helps to break up the speech stream into smaller units. Stressed syllables are either louder, higher, and/or longer than unstressed syllables and more resistant to reduction processes. Of all the types that stress comes in, we will focus on word stress here (as opposed to higher-level stress such as phrasal or sentence stress). In the following we will have a look at some basic concepts of stress (Section 2), before we come to a case study of Modern Greek stress (Section 3) and how it can be acquired in computer simulations (Section 4). Modern Greek stress is intriguing because of its complexity: there are words with final stress, others with penultimate stress and yet others with antepenultimate stress, irrespective of the phonological structure of the words. On top of that, stress can shift to the right depending on the case (at least in nouns). These shifts are not always realized, adding to the complex pattern. This makes the learnability of this stress system, that is, a theoretical approach on how this pattern could in principle be acquired, particularly interesting.

Some words on the terminology: the notions ‘stress’ and ‘accent’ are often used interchangeably in the literature, and sometimes used as discriminating between different phonetic aspects. ‘Accent’ is in general used as a slightly more general term and often refers to sentence accent. Here, ‘accent’ will be used to distinguish between the Ancient Greek pitch accent system versus the Modern Greek dynamic stress system.

2. Phonetic and Phonological Aspects of Stress

2.1 Phonetic Aspects

Stress manifests phonetically with one or more correlates. Auditorily or perceptually speaking these are length, loudness, pitch and vowel quality. Acoustically speaking, the correlates are reflected in duration, intensity or amplitude, fundamental frequency and spectral structure (e.g. Dogil and Williams, 1999; cf. Fry, 1955; Lehiste, 1970). The use of these correlates is very language-specific. English, for one, is said to have pitch as a stress correlate, combined with a preservation of vowel quality in unstressed syllables and a strong tendency to reduce vowels in unstressed syllables. In Modern Greek, the language of our case study, stress is characterized by a combination of intensity and duration (McKeever Dauer, 1980; Botinis, 1989). In German, duration/length seems to be the main cue for word stress (Dogil and Williams, 1999; Jessen and Marasek, 1995).

All of the above correlates are relative, and all of them can express something else than stress. For instance, the duration of a vowel can be a sign for stress in that stressed syllables are often longer in duration than unstressed ones, but it can also be a sign of distinctive vowel length. Pitch can, other than stress, express tone, or prosody on the sentence level. Despite this inconclusiveness, in a lot of languages speakers usually agree in their judgement about which syllable in a word is the most prominent. We now turn to a brief overview over some functional aspects of stress.

2.2 Functional Aspects: Predictable Versus Lexical Stress

Stress is said to have several different functions. Depending on the language, the assignment of stress can be rather predictable (e.g. in languages where the beginning or the end of words is canonically stressed) or it can be unpredictable in the sense that it needs to be learned for each word (so-called lexical stress¹). In the former case, stress could serve as a means to signal a boundary (see Trubetzkoy, 1939; Hyman, 1977; Karvonen, 2008 for the idea of stress as

a demarcative function). In the latter case, stress can carry information about the morphological or lexical make-up of a word. Predictable stress systems are also referred to as *fixed* stress systems, whereas lexical stress systems are also referred to as *free* stress systems.

Languages with predictable stress assign it by phonological principles, for instance to a certain position in the word. Consider Finnish, where words are stressed initially as shown in (1) (indicated by an acute accent; examples are taken from Karvonen, 2008):

- (1) Finnish:
 (a) *vápaá* 'free'
 (b) *hélsinki* 'Helsinki'
 (c) *érgonómia* 'ergonomics'

In Turkish, stress is for the most part final. No matter how many suffixes are attached, stress is usually on the ultimate syllable as shown in (2) (examples are from Sezer, 1983):

- (2) Turkish:
 (a) *tanı-dík* 'acquaintance'
 (b) *tanı-dík-lár* 'acquaintances'
 (c) *tanı-dík-lar-ím* 'my acquaintances'

Stress is in these cases edge-oriented, either towards the left or right edge of words. This does not have to be the initial or final syllable; it can also be near an edge.² Regular penultimate stress is common (e.g. in Polish; for example, Comrie, 1976; Hayes, 1995a and references therein); antepenultimate (e.g. in Macedonian; Comrie, 1976) and even postinitial stress is found (e.g. in Dakota; Chambers, 1978). Other phonological factors that can play a role in stress assignment are the complexity of the syllable (*heavy* syllables often attract stress in contrast to *light* syllables; see Section 2.4).

Languages with largely unpredictable stress patterns are, among others, Russian (Halle, 1997; Melvold, 1990; Revithiadou, 1999)³, Bulgarian (Dogil et al., 1999: 843), and Modern Greek (e.g. Philippaki-Warburton, 1970, 1976). In Russian, stems and affixes can be inherently ('lexically') marked for stress, and words with marked stems have stress on the stem-syllable (3a and b), whereas words with unmarked stems have stress on the affix (if that is lexically marked; 3c) or on the initial syllable (in case no morpheme is marked for stress; 3d).

- (3) Russian:
 (a) *rabót+á* = *rabóta* 'work-Nom.Sg.'
 (b) *rabót+y* = *rabóty* 'work-Nom.Pl.'

- (c) borod+á = *borodá* 'beard-Nom.Sg'
- (d) borod+y = *bórody* 'beard-Nom.Pl'

Modern Greek, our case study and described in more detail in Section 3, behaves quite similar, except that it has a 'trisyllabic window', that is, stress can only occur on one of the final three syllables.

In general, the distinction between predictable and unpredictable stress is continuous rather than absolute: languages with predominantly predictable stress can have exceptions and languages with unpredictable stress can have some phonological restrictions or a preference for a certain pattern. Languages like English, Dutch or German appear to have stress based to a great extent on phonological properties (van der Hulst, 1999, and references therein), but morphology and the lexicon play a role, too. In English, verbs and nouns are sometimes only distinguished by their different stresses: *récord* (noun) versus *recórd* (verb). Since the segmental make-up of verb and noun are the same, the difference in stress has to be lexical rather than phonological. The morphological influence on stress in English furthermore shows in words such as *recordability*, where the derivational suffix attracts stress away from the root.

2.3 Examples of Languages without Word Stress

Most languages make use of some sort of stress on the word level, however, a few languages appear to lack it altogether (e.g. Betawi Malay, Roosman, 2006; Javanese Indonesian, Goedemans and van Zanten, 2007; Sri Lankan Malay, Nordhoff, 2009; Apoussidou and Nordhoff, 2010), or at least have some content words that lack it (Cayuga and Seneca, Chafe, 1977; Central Sierra Miwok, Freeland, 1951; Japanese, Poser, 1990; Kinyambo, Bickmore, 1989, 1992; Yupik Eskimo, Krauss, 1985). It is debatable whether French (Dell, 1984) has word stress or not, or whether it only has stress on the phrasal level (Hayes, 1995a). Grammatical words are frequently unstressed. The occurrence of these languages suggests that the principle of *culminativity* (that each word or phrase should have one and only one main stress; Liberman and Prince, 1977), once seen as an unviolable universal, is either not universal or not unviolable. At least since the advent of Optimality Theory (Prince and Smolensky, 2004 [1993]) the latter has become an option (e.g. Alderete, 1999).

2.4 Phonological Aspects of Stress

In systems with only one stress per word that furthermore falls on one of the two possible edges, it suffices to designate the edge as a phonological rule or

constraint. These systems are called *unbounded* (Halle and Vergnaud, 1987). *Bounded* stress systems, for instance patterns with stress on the second or third syllable from the edge, and patterns with iterative stress (i.e. patterns that have secondary stresses besides the main stress of a word) could be accounted for by counting syllables; this, however, is a rather inelegant approach. Given the binary nature of stress patterns, for which counting is not necessary, one came up with templates called *feet* that describe the observed patterns. An overview over the initiation of the concept ‘foot’ is given in Hayes (1995a: 40). Hayes furthermore established the foot inventory that is still the basis for most work on metrical issues. The notion of ‘feet’ stems, among others, from grid theory (Prince, 1983; Selkirk, 1984b), where stressed syllables were displayed as grid marks above a segmental tier, and unstressed ones as dots. This representation, exemplified in (4), is even reminiscent of a stylized foot (here with the leg on the left side and the toes pointing to the right):

- (4) *
 (* .)
 mo.ra

Feet were invoked to explain the often alternating behaviour of stressed and unstressed syllables, and the location of stress within a word, but they are not audible in the sense that stress is. They are often binary (i.e. consisting of two syllables), as, for example, languages such as Pintupi (Hansen and Hansen, 1969) in (5) and Finnish in (6) suggest. In these languages, secondary stresses are assigned in an alternating manner to every other syllable next to the main stress:

- (5) Pintupi:
múŋu → (múŋu) ‘orphan’
ŋáلكunìnpa → (ŋáلكu)(nìnpa) ‘eating’
tʃámulìmpatʃùŋku → (tʃámu)(lìmpa)(tʃùŋku) ‘our relation’
- (6) Finnish:
érgonòmia → (érgo)(nòmi)a ‘ergonomics’

Languages like Cayuvava (as in Hayes, 1981, 1995a) can be explained with trisyllabic feet: for example, (*jìhira*)(*ríama*) ‘I must do’, *ma*(*ràhaha*)(*éiki*) ‘their blankets’, *iki*(*tàpare*)(*répeha*) ‘the water is clean’. In other languages such as Modern Greek, monosyllabic feet can play a role (as we will see later).

Crucial to the notion of metrical feet is the division into a strong part (the *head*) and a weak part. Left-headed feet are called trochees, right-headed feet are called iambs. As indicated in (5) and (6), Pintupi and Finnish can be

analysed as having trochees. Pacific Yupik (as in Hayes, 1995a: 335 and references therein) can be analysed as having iambic feet as shown in (7):

- (7) Pacific Yupik:
atáka → (atá)ka 'my father'
akútamák → (akú)(tamák) 'kind of food-abl.sg.'

Iambic languages tend to make a distinction between heavy and light syllables (unstressed syllables are short/shortened and stressed syllables are lengthened/long), whereas trochaic languages do not. This was formulated in an *Iambic-Trochaic Law* (going back to a perceptual experiment by Bolton, 1894): 'elements contrasting in intensity naturally form groupings with initial prominence', and 'elements contrasting in duration naturally form groupings with final prominence' (phrasing taken from Hayes, 1995a: 80). Across languages, this seems to be more of a tendency than a law (see Kager, 1993).

Speaking of heavy and light syllables, they require some explanation. There are languages that do not make a distinction between heavy and light syllables irrespective of their syllable structure. Some languages distinguish between open syllables with a short vowel that are light, and open syllables with a long vowel that are heavy. In these languages, the heavy syllables tend to attract stress. Adding to these are languages that count syllables with closed syllables as heavy (e.g. Latin, Mester, 1994; German, Giegerich, 1985) or light. Languages that do not have a vowel length distinction but count closed syllables as heavy (e.g. Spanish, Roca, 1999) are apparently rare. Modern Greek has closed syllables (i.e. syllables ending in a consonant), but no vowel length distinction. Stress in this case is assigned irrespective of whether a syllable is closed or not.

These are just the very basic principles of stress. As for the acquisition of stress, a learner of a language (at least an L1 learner) has to find out whether his/her language is predominantly predictable, and if so, which phonological grammar renders the correct stress pattern, and if not, what needs to be represented in the lexicon. Let us now turn to a more detailed discussion of stress in Modern Greek, before we move on to the computational treatment.

3. Modern Greek

Modern Greek word stress, especially in nouns, is complex and involves an interaction of the lexicon and the phonology of the language. Modern Greek (henceforth simply Greek, unless it is contrasted with Ancient Greek) has lexical stress, meaning that the morphemes of a word can be (and often are) underlyingly specified for stress. Stress can occur on any of the final three syllables (depending on the word), and in the genitive case stress can shift to

the right. It is phonologically restricted in the sense that stress cannot occur farther to the left than on the antepenultimate syllable. The shift in the genitive case is an artefact of Ancient Greek, where it was phonologically conditioned. Nowadays, the shift is lexical and not transparent anymore: speakers use it when using a more formal speech style, but some words undergo it more easily than others. The genitive case is furthermore not frequently used, and speakers reveal insecurities when prompted to produce the genitive stress pattern for certain words. This can be due to the infrequency of the genitive case, that is, the speakers were not exposed to the genitive case enough to fully acquire the pattern, but it can also be due to the fact that the stress shift in the genitive case is an artefact of Ancient Greek and is in conflict with the contemporary phonology of Greek.

This chapter will give a computational learnability account of word stress in Modern Greek nouns, showing how the lexical word stress can be acquired using allomorphy, and furthermore showing how some of the stress patterns occurring in Modern Greek could be more difficult to acquire than others due to both the restrictions of the grammar and to the infrequency of the genitive case.

3.1 The Data

Stress in Greek is lexical in the sense that morphemes can be underlyingly specified for it. Any of the last three syllables of a word can be stressed (see (8); examples are taken from Holton et al., 1997/2004; dots indicate a syllable boundary and acute accents indicate stress; citation form is the nominative singular, if not specified otherwise):

(8) Possible stress positions in Modern Greek

antepenultimate	penultimate	ultimate
<i>thá.la.sa</i> 'sea'	<i>kó.ri</i> 'daughter'	<i>a.go.rá</i> 'market place'
<i>va.sí.li.sa</i> 'queen'	<i>el.pí.tha</i> 'hope'	<i>u.ra.nós</i> 'sky'
<i>la.ví.rin.thos</i> 'labyrinth'	<i>fi.ga.té.ra</i> 'daughter'	<i>a.del.fós</i> 'brother'
	<i>e.fi.me.rí.tha</i> 'newspaper'	

In the genitive case, stress can shift from the antepenultimate of the nominative case to the penultimate syllable (9a), or to the ultimate syllable (9b), or from the penultimate to the ultimate syllable (9c). In other cases, stress does not shift (9d).

- (9) a. *la.ví.rin.thos* ~ *la.vi.rín.thon*
 b. *thá.la.sa* ~ *tha.la.són*; *va.sí.li.sa* ~ *va.si.li.són*

- c. *do.má.ta* ~ *do.ma.tón*
- d. *el.pí.tha* ~ *el.pí.thon*

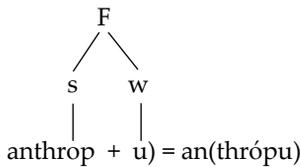
In *any* case, stress cannot occur left of the antepenultimate syllable.⁴

3.2 Previous Analyses

There is consensus in the literature that stress in Greek is for the most part lexical. A phonological default applies only when no morpheme of the word is specified for stress; in this case, stress is assigned to the antepenultimate syllable if there is one (Philippaki-Warburton, 1970, 1976; Ralli, 1988; Malikouti-Drachman and Drachman, 1989; Ralli and Touratzidis, 1992; Drachman and Malikouti-Drachman, 1996; Revithiadou, 1999).⁵ Both roots and suffixes can be specified for stress.

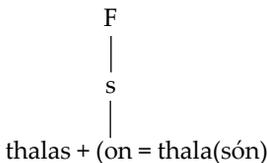
Revithiadou (1999), for instance, proposed that the stress shift of the genitive singular suffix *-u* is caused by an underlying specification in form of a weak foot part. The preceding root gets stressed on its final syllable because the grammar of the language builds a trochaic foot based on the specification of the suffix. While a word like *ánthrop+os* ‘human’ is stressed by the phonology, that is, (ánthro)pos, the genitive singular case *-u* comes with a lexical foot part that the phonology incorporates into a trochaic foot to form an(thrópu) as shown in (10) (‘s’ stands for the strong part of a foot; ‘w’ for the weak part):

(10) Pre-stressing foot structure:



The genitive plural suffix *-on* is in Revithiadou’s (1999) analysis stressed itself, that is, equipped with a strong foot part underlyingly. This results in a monosyllabic foot for the suffix as shown in (11):

(11) Self-stressing foot structure:



Both suffixes can attract stress from the antepenultimate syllable only if adhered to an underlyingly unstressed root. As the data in (12) and (13) show, stress can shift from the antepenult to the penult or ultima, where the suffix *-on* in some cases surfaces as pre-stressing and sometimes as stressed itself (and not only when occurring with roots that are unstressable).

- (12) Stress shift from the antepenultimate to the penultimate syllable:

méthodos ~ *methódon* 'method'
ánthropos ~ *anthrópon* 'human'

- (13) Stress shift from the antepenultimate to the ultimate syllable:

thálasa ~ *thalasón* 'sea'
trápeza ~ *trapezón* 'bank'

This already indicates that there is some kind of allomorphy involved. As the data in (14) furthermore show, stress can also shift from the penult to the ultima:

- (14) Stress shift from the penultimate to the ultimate syllable:

domáta ~ *domatón* 'tomato'
turístas ~ *turistón* 'tourist'

According to Revithiadou's analysis, the suffix should only be able to attract stress away from the root if the root is underlyingly unstressed. If a word surfaces as stressed on the penultimate syllable in the nominative case, then it has to be underlyingly stressed on that syllable, because otherwise it should be stressed by default on the antepenultimate (if the root is disyllabic or longer). If the root is underlyingly stressed, then the inflectional suffix should not be able to attract stress away from the root due to the ranking of root stress over suffix stress. This kind of stress shift cannot be explained by underlying foot structure as in Revithiadou's analysis (nor by Apoussidou's 2007 account for that matter).

An alternative would be allomorphy: in the penult-to-the-ultima case; one could assume either stressless and pre-stressing allomorphs for the non-genitive suffixes, assigning penultimate stress, or one could assume that the root has two allomorphs; one unstressed and one stressed on the final syllable, that is, *méthod-* ~ *methód-*. The latter might be less costly for a learner of the language since it is a straightforward interpretation of the input. Moreover, allomorphy exists anyway in the language. Drachman et al. (1995) give an account how allomorphs of Greek suffixes are chosen on the basis of their own shape and the shape of the root/stem they adhere to. For instance, action nominals can be formed with the suffixes *-imo* and *-ma* as given in (15). Monosyllabic stems take

the disyllabic *-imo*, while polysyllabic stems take the monosyllabic *-ma* (examples taken from Drachman et al., 1995):

(15) Action nominals

vréko 'I wet' > *vréksimo* 'wetting'
skupiso 'I sweep' > *skúpizma* 'sweeping'

However, Drachman et al. (1995) are suspicious of suffixal alternants that are segmentally identical but behave differently with respect to stress, as in the deverbal agent nominalizer *-tis*. This suffix surfaces as stressed or unstressed depending on the length (in terms of syllable count) of the preceding root: *kléf-tis* 'thief' versus *skupis-tís* 'sweeper'. Instead of allomorphy they propose catalexis for the behaviour of suffixes like *-tis*, to account for why the unstressed variant adheres to monosyllabic stems and the stressed variant adheres to polysyllabic stems. While they give an elegant account for the distribution of (other) allomorphs on prosodic grounds in Modern Greek, I am less suspicious of unstressed *-tis*/stressed *-tís* as allomorphy, because the catalectic analysis does not apply in a similar case, the genitive plural suffix *-on*. This suffix behaves like the *-tis/-tís* case in that it has stressed, pre-stressing and unstressed alternants, but it alternates irrespective of the prosodic shape of the root it adheres to (as I will show below). I would like to argue that roots and suffixes can simply have unstressed and stressed alternants, and they adhere to each other by convention, not necessarily out of prosodic reasons. The reason for this assumption is that the stress-shifting pattern caused by some affixes usually belongs to a more formal register stemming from Ancient Greek (Holton et al., 1997/2004) and is 'learned'. I therefore take that the stress-shifting pattern is highly lexicalized and possibly not part of the natural language acquisition process.

Since the examples show that (i) stress is lexical in Greek; (ii) allomorphy is needed in one way or another anyway (*méthodos* ~ *methódon* and *thálasa* ~ *thalasón*) and (iii) underlying foot structure cannot account for all the cases of stress shift (*domáta* ~ *domatón*), I would like to propose that Greek makes use of underlying stress, but that the stress shifts are encoded as allomorphy, as exemplified in (16) (similar to Ralli and Touratzidis' 1992 approach).

(16) Simplified underlying representations:

méthodos ~ *methódon*; root allomorphs: |méthod-| and |methód-|
thálasa ~ *thalasón*; root allomorphs: |thalas-| and |thálas-|
uranós ~ *uranón*; suffix allomorphs: |-ós| and |-ón|

One more note is in order. The allomorphic behaviour of the affixes cannot be attributed to the declension class or gender of the nouns they belong to. For one, *-on* is the genitive plural marker for practically all nouns (except for the ones

that only occur in the singular). The distribution of stressed/unstressed/pre-stressing *-on* is rather arbitrary: whether or not a noun takes the stressed or unstressed version does not depend on gender or declension class or prosody (classification basically taken from Holton et al., 2004; I left out imparisyllabic nouns), but on its phonological structure in Ancient Greek.

Drachman et al. (1995), Revithiadou (1999) and Apoussidou (2007) assume underlying foot structure as specification (Inkelas, 1994), to account for stress-shifting pattern in nouns caused by, for example, the genitive plural suffix *-on*. I will give an alternative account here where morphemes are underlyingly stressed in the language, but where there is no need to assume underlying foot structure or pre- and post-stressing mechanisms. Instead, there are only two straightforward underlying specifications: stressed and unstressed. The complex pattern in the language comes about with allomorphy: roots and affixes have underlyingly stressed and unstressed variants. For acquiring the stress patterns, no intricate underlying representations have to be acquired other than what can be observed in the data.

3.3 The Shift: A Relic

The stress pattern of Modern Greek is complex because it developed out of the Ancient Greek (AG) pitch accent. This accent was lexical as well: any of the last three syllables of a word could bear one. But the shift of the accent within the paradigm (the ‘recessive accent’) was predictable: stress would shift if the final syllable was ‘heavy’ (see Steriade, 1988; Kiparsky, 2003). Final syllables were heavy in AG if they contained either a long vowel (AG had phonemic vowel length) or ended in a consonant cluster. The genitive plural suffix *-on* contained a long [o:], written as ‘ω’. The trisyllabic window in AG was therefore a *trimoraic* one (cf. Smith and Apoussidou in prep.). This explains why the accent was attracted away from the antepenultimate syllable in cases like *lavírinthos* ~ *lavírinthon*. In *lavírinthos*, the final syllable is light (*-os* contained a short vowel: *laví_μri_μntho_μs*) and the accent is on the antepenultimate syllable because this syllable happens to contain the third mora from the end. In *lavírinthon*, the final syllable is heavy (because the /o/ in *-on* is long: *laví_μri_μntho_μn*), causing the stress to shift a syllable to the right. An accent on the antepenultimate syllable would violate the trimoraic window: **laví_μri_μntho_μn*.

Newton (1972: 12) suggests that vowel length is still preserved in the modern language underlyingly, arguing that under this assumption, the stress shift remains predictable. However, Philippaki-Warburton (1976) mentions that speakers do not use the shifted forms as much anymore, indicating that speakers are not aware of the Ancient Greek distinction. I argue (in line with Philippaki-Warburton, 1976) that speakers of the modern language are not

aware of the underlying phonemic vowel length proposed by Newton and that the stress shift is just lexically conditioned, because (i) there is no evidence of phonemic vowel length in the phonetic signal, so the question is how speakers of Greek come to this knowledge and (ii) nowadays, stress in at least more informal speech does not shift as often anymore (see also Holton et al., 1997/2004), which is not explained under the assumption that underlying vowel length is somehow transparent.

It thus appears that the loss of the vowel length distinction pulls the rug out from under the phonological condition that caused the stress shift.⁶ And since the language already made use of lexical stress anyway, I argue that at least in the nominal paradigm, stress is lexical all the way.

3.4 Frequency Patterns

As of now, not much quantitative data is available about how many words exist with the different patterns, and how many of them shift. Revithiadou (1999) compiled a small corpus of 16,000 nouns in *-os* and *-a*, of which she classified 67.5% as accented, 18.5% as unaccentable roots (indicating that these words do not shift), and 10.2% as unmarked ones (indicating that these shift in the genitive case). Protopapas (2006) indicates that penultimate stress is the most frequent (but not overwhelmingly frequent) pattern in Modern Greek. However, his count is based on words with two or more syllables, and it is not clear whether antepenultimate stress is simply less frequent due to the fact that disyllabic words cannot be stressed pre-penultimately. Protopapas et al. (2006) tested whether the stress diacritic in written Modern Greek is necessary for correctly reading a word or whether children have alternative strategies for stressing. Again, penultimate position seems to be the most favoured position for stress in Modern Greek, but there is no information about stress-shifting patterns. Sims (in press) examined Modern Greek speakers in their use of periphrastic constructions to avoid stress-shifting patterns in Modern Greek, but her sample of nouns does not indicate a percentage of stress-shifting versus non-shifting patterns. None of these studies give a count of the different patterns in terms of possible variation as mentioned by Holton et al. (1997/2004) or Philippaki-Warbuton (1976). The goal of this chapter is not a quantitative survey of the occurring variation; but in the following, I will give a sample of it, drawn from the Hellenic National Corpus (HNC; www.hnc.ilsp.gr/en/default.asp) of written texts. The examples are composed of words with three to five syllables and compared with Holton et al. (1997/2004). Words with final stress in the nominative case are disregarded for the moment, since they never show variation. According to Holton et al. (1997/2004; henceforth HMP), the words for 'sea' and 'tomato' should always shift in the genitive (HMP: 57, indicated

here by i); the words for ‘almond’ and ‘butter’ should show variation (HMP: 64, indicated here by ii), and ‘headache’ should not shift (HMP: 52, indicated here by iii) as shown in (17).

(17) Sample of frequencies (empty cells rendered a count of zero):

gloss	nominative	genitive (shift)	genitive (no shift)	ratio
‘sea’	<i>thálasa</i> (2831)	<i>thalasón</i> (94)	i	30:1
‘tomato’	<i>domáta</i> (105)	<i>domatón</i> (1)	i	105:1
‘almond’	<i>amígdalo</i> (16)	<i>amigdálon</i> (3)	ii	5:1
‘butter’	<i>vútiro</i> (99)	ii	<i>vútiron</i> (1)	99:1
‘sheep’	<i>próvato</i> (119)	<i>prováton</i> (50)	<i>próvaton</i> (2)	2:1 60:1
‘man’	<i>ánthropos</i> (6091)	<i>anthrópon</i> (4281)	<i>ánthronon</i> (11)	1,5:1 554:1
‘headache’	<i>ponokéfalos</i> (177)	<i>ponokefálon</i> (3)	iii	59:1

Some variation was found in the genitive singular instead of the genitive plural (not listed here). The genitive case is rather rare, with the exception of high-frequency nouns such as *ánthronon* and *thalasón*. There is also no clear pattern as of how frequent one genitive form is over the other, if there is variation. In general, the shifting forms are more frequent than the non-shifting forms (with the exception of one occurrence of *vútiron*). Since the HNC is a corpus of written material only, this rather conservative pattern might be reflecting a more formal register than would be used in spoken language. It suffices to indicate that indeed there exists variation in the stress patterns; if anything, the variation would be bigger in the spoken language. This variation is apparently acquired by some of the speakers at least, and its learnability will be accounted for here.

3.5 Harmonic Allomorphy Instead

I argue that the stress pattern is highly lexical and that the use of allomorphs can account for both the shifting and the non-shifting patterns, but that some of the shifting patterns are more difficult to acquire than others. This can be formalized by adopting a framework such as Harmonic Grammar (Legendre, Miyata and Smolensky, 1990a; Smolensky and Legendre, 2006), differing from Optimality Theory (henceforth OT) in that constraints can gang up to choose a candidate not preferred by a strict ranking hierarchy.

Since the learner cannot know beforehand whether s/he is learning a language with lexically or grammatically assigned stress, we want to leave the choice up to the learner for which kind of stress s/he finds evidence. The simplified assumption is that learners are provided with the meaning of

the phonetic (or overt) form they hear, and have to create the corresponding phonological surface structure and lexical (underlying) representation. There are hence four levels of representation involved: meaning (word meaning and syntactic information such as gender, number or case), the underlying representation (the root and suffix morphemes with or without stress marks), the phonological surface structure (containing feet and stress), and the phonetic overt form (containing only stress).

Determining the underlying representation for an incoming form works as follows. Learners adopt underlying representations based on what can be observed on the overt level (cf. Smolensky, 1996 and references therein). A word that is observed in the nominative singular as stressed on the root, for example, *thálasa*, leads to the creation of an entry with a stressed root morpheme |thálas-| and an entry with an unstressed suffix |-a|, in form of *lexical constraints* (cf. Boersma, 2001; Apoussidou, 2007) connecting the meaning and syntactic information of the morphemes with their phonological representation. Is the word subsequently observed in the genitive plural form as *thalasón*, an unstressed allomorph for the root is created, |thalas-|, as well as a stressed suffix morpheme |-ón|.

In the following, I very roughly outline a Harmonic Grammar and apply it to a ‘grammar’ containing lexical constraints. The lexical constraints stem from what can be observed in the overt form (displayed in square brackets ‘[]’): a stress on the root in the overt form gives rise to the creation of a lexical constraint linking the meaning (in angle brackets ‘<>’) of this root to an underlying representation (in pipes ‘| ’) of this morpheme containing the segments and a mark for the stress on the designated syllable.⁷ For instance, the root ‘sea’ sometimes appears as stressed and sometimes as unstressed: stressed if inflected with the nominative case suffix *-a*, and unstressed if inflected with the genitive plural suffix *-ón*. Likewise, a constraint for unstressed *-a* and stressed *-ón* are created. The meaning is used to index the underlying forms as shown in (18):

(18) Creation of lexical constraints

given:	created:
‘sea-Nom.Sg.F’ [thálasa]	→ <sea> thálas- ; <Nom.Sg.F> -a
‘sea-Gen.Pl’ [thalasón]	→ <sea> thalas- ; <Gen.Pl> -ón

For the genitive plural suffix, a lexical constraint with an unmarked *-on* will be created for words such as *prováton* ‘sheep-Gen.Pl’, and so on and so forth. The learner will also posit surface structure containing feet, representing the phonology of the language. Because we want to focus on the creation of the lexicon, we ignore the phonology for the moment (but will include it in the simulations).

In Harmonic Grammar (HG), constraints are not strictly ranked as in OT, but are assigned numerical weights, as indicated by the numbers on top of the

constraint names in Tableau 5.1. Each violation of a constraint is assigned a negative value of -1. The harmonic score (the number that determines the winning candidate) is calculated by multiplying the violation value with the constraint weight, and taking the sum of all the products for a candidate. The harmonic value closest to zero (here: -2.5) determines the winning candidate; in the case of Tableau 5.1 it is the candidate without any lexical specification for stress, which is preferred by the constraint without underlying lexical specifications for the root 'sea'. For ease of exposition, I only use the lexical constraints evaluating the underlying representations, leaving out the surface structures and overt forms.

Tableau 5.1 Determining the underlying representation for morphemes in HG: [thalas+a]

	3	2.5	1	
<sea+Nom.Sg.F>	<sea> thalas-	<sea> thálas-	<Nom.Sg> -a	
☞ thalas+a		-1		-2.5
thálas+a	-1			-3

Under the assumed weighting of constraints in Tableau 5.1, the word *thálasa* in the nominative singular is analysed as underlyingly unmarked for stress. In combination with the genitive plural suffix, different constraints and candidates come into play: combinations of unstressed/stressed roots and unstressed/stressed suffixes (because that is what can be observed for Greek). Tableau 5.2 shows how the candidate is evaluated as the most harmonic one that has an underlyingly unstressed root and the underlyingly stressed allomorph of the genitive plural suffix, *-ón*: the candidate violates the lexical constraint for a stressed root and the constraint for an unstressed suffix, but in sum provides the harmonic score closest to zero.

Tableau 5.2 Choosing the stressed allomorph of <Gen.Pl>: [thalasón] → [thalas+ón]

	3	2.5	1.1	1	
<sea+Gen.Pl>	<sea> thalas-	<sea> thálas-	<Gen.Pl> -ón	<Gen.Pl> -on	
thalas+on		-1	-1		-3.6
☞ thálas+ón		-1		-1	-3.5
thálas+on	-1		-1		-4.1
thálas+ón	-1			-1	-4

In this way, the shift from the antepenultimate syllable to the ultimate syllable can be accounted for without assuming underlying foot structures that would require ad hoc assumptions about how to acquire them.

Next to the shift from the antepenultimate to the ultimate syllable exists the shift from the antepenult to the penult in Greek. It turns out that this shift is less easy to accomplish, given the assumed weighting of the constraints in Tableau 5.2. Tableau 5.3 shows that the actual winner is the candidate with a specification on both the root and the suffix; but this yields underlying stress on the incorrect syllable. If this candidate is chosen as the underlying form, a surfacing stress on the antepenultimate or ultimate syllable is likely. There are two candidates that have a harmony score close to the winning candidate of Tableau 5.3, |próvat+on| and |provát+ón|. The latter candidate has the root allomorph that could result in the aspirated shift of stress if the weights for <sheep> |próvat-| and <sheep> |provát-| are switched. However, this would lead to another winning candidate in the nominative form as well.

Tableau 5.3 Shift from antepenultimate to penultimate syllable? *|próvat+ón| /pro(váton)/:

	1.5	1.4	1.1	1	
<sheep+Gen.Pl>	<sheep> próvat-	<sheep> provát-	<Gen.Pl> -ón	<Gen.Pl> -on	
provat+on	-1	-1	-1		-4
provat+ón	-1	-1		-1	-3.9
próvat+on		-1	-1		-2.5
☞ próvat+ón		-1		-1	-2.4
provát+on	-1		-1		-2.7
provát+ón	-1			-1	-2.5

There is no weighting of the assumed constraints that both accounts for Tableau 5.2 and Tableau 5.3. If the lexical constraint <sheep> |provát-| outweighs the constraint <sheep> |próvat-|, then the root allomorph with final stress would win, something that under a faithful evaluation would result in the wrong surface structure with penultimate stress in the nominative case as shown in Tableau 5.4.

Note that we are still only talking about underlying representations here. There are two ways that could make an antepenultimately stressed representation win in the nominative case and a penultimately stressed representation in the genitive case. One is that unfaithful surface structures assign stress

Tableau 5.4 Re-weighting the constraints

	1.5	1.4	1.1	
<sheep+Nom.Sg>	<sheep> provát-	<sheep> próvat-	<Nom.Sg> -o	
provat+o	-1	-1		-2.9
próvat+o	-1			-1.5
☞ provát+o		-1		-1.4

differently than underlyingly present. The other is adopting a stochastic evaluation of the constraint weights (similar to stochastic OT): the candidate with the finally stressed root |provát+ón| could still win in Tableau 5.3 in some cases if one assumes noise that is added to each evaluation. This slight change in the weighting can result in the selection of a different candidate. Both options will be included in the simulations below. The prediction is that the learners exposed to lexical stress on different positions within a word, but without stress shifts in the genitive case, will learn the pattern correctly, while learners exposed to a stress pattern with a shift in the genitive case will acquire the shift only partially.

4. Simulating the Acquisition Process

4.1 Computational Studies on the Learnability of Stress Patterns

There has been a considerable amount of research on the computational learnability of stress. Most studies focused on the learning of the grammatical part (e.g. Dresher and Kaye, 1990; Clark and Roberts, 1993; Daelemans et al., 1994; Gupta and Touretzky, 1994; Tesar, 1998; Tesar and Smolensky, 2000; Heinz, 2006, 2009; Hayes and Wilson, 2008), especially on the acquisition of quantity-sensitive or -insensitive systems. More recently the problem of learning on how the respective underlying material could be acquired has been tackled (e.g. Tesar, 2004, 2009; Alderete et al., 2005; Jarosz, 2006). Most of these approaches assume *offline* or *batch* learning, that is, learning from the complete set of data presented to the learner at the same time. Since it is more likely that infants process input *online*, with one data item at a time, we need an approach capable of that. Further, since we are dealing with a language here that has unpredictable, that is, lexical stress, in combination with a phonological restriction, the focus here will be on the acquisition of underlying material in line with the acquisition of the grammatical regularities. For that,

Apoussidou's (2007) approach of learning underlying forms is adopted, which is capable of handling several levels of hidden structure.

4.2 Tackling the Acquisition of Hidden Structures

As outlined, the stress pattern in Greek is complex, but can be broken down into three basic patterns: (i) the phonological default and trisyllabic restriction assigning stress on the antepenultimate (if there is one); (ii) lexical stress on different morphemes resulting in stress on the antepenultimate, penultimate and ultimate syllable; (iii) like (ii), but additionally a stress shift in the genitive case. The patterns co-occur and it is likely that most speakers of Greek use all three patterns, resulting in a hybrid stress system. To test the learnability of the different stress patterns in Greek, I set up different groups of virtual learners that are trained on the three basic stress patterns and on two hybrid patterns. Learners are considered as being successful if they render the pattern they have heard correct, that is, if they produce the words they were trained on with the same stress pattern. In addition to that they will be tested on how well they generalize to words they have not heard. All learners are noisy Harmonic-Grammar learners (Boersma and Pater, 2008; see also the computer settings in Section 4.6) and have the same structural and faithfulness constraints. They differ in the lexical constraints depending on which words they are exposed to. The weighting of these constraints comprises the lexicon. The structural constraints are given a head start over the lexical and faithfulness constraints (cf. Jesney and Tessier, 2008), simulating the creation of lexical and faithfulness constraints that the learners should in principle invoke themselves (in lack of a lexical and faithfulness constraint induction mechanism).

4.3 The Training Sets

I tested five different data sets listed in (19), meaning that different groups of virtual learners are trained on different stress patterns, all including forms in the nominative and genitive case with a ratio of 3:1 (i.e. for each occurrence of the genitive case there are three occurrences of the corresponding nominative case form).

- (19) The training sets:
- (a) No shift/default
 - (b) No shift/lexical
 - (c) Always shift
 - (d) Sometimes shift
 - (e) Some shift

The different training sets are idealized learning conditions; in real life, speakers of Greek are exposed to different frequencies of all these patterns. The pattern in (19a) is a test of the phonological default only, whereas pattern (19b) comes close to the informal speech style without stress shift. Pattern (19c) comes close to the formal speech style where stress always shifts. The hybrid patterns of (19d) and (19e) come closest to a real-life scenario: ‘sometimes shift’ means there are some words that have varying stress in the genitive case, and ‘some shift’ means there are certain words that always shift in the genitive case and others that never shift.

4.3.1 No Shift and a Fixed Position for Stress

The ‘no shift/default’ training set of the list in (19) tests whether the virtual learners acquire the phonological default (stress on the antepenultimate syllable if there is one) and the trisyllabic window restriction (stress not further to the left than the antepenultimate syllable). It contains four 2–5-syllable words that have stress on the penult (the 2-syllable words) or on the antepenult, and stress does not shift in the genitive case (20).

The training items consist of the meaning paired with the phonetic form. Our virtual learners are fed rather these pairs of overt (phonetic) forms and their corresponding meaning than just the phonetic form to enable them to distinguish between words with lexically and words with phonologically assigned stress. The overt forms contain segmental information and stress (but no information on foot structure), and the meaning contains the lexical and syntactical information. This means in effect that the virtual learners already know that the words contain a morpheme boundary and also where that boundary is; two things that a learner under more natural conditions is not provided with.

(20) The ‘no shift/default’ training set:⁸

Nominative singular		Genitive plural	
<tree+Nom.Sg.N>	[déndro]	<tree+Gen.Pl>	[déndron]
<sheep+Nom.Sg.N>	[próvato]	<sheep+Gen.Pl>	[próvaton]
<almond+Nom.Sg.N>	[amígdalo]	<almond+Gen.Pl>	[amígdalon]
<headache+Nom.Sg.M>	[ponokéfalos]	<headache+Gen.Pl>	[ponokéfalón]

What the virtual learners had to come up with themselves were the surface structures (generated by the grammar of the learners; here: feet and stresses) and the underlying representations (generated by the lexicon of the learners; here: underlying stress versus no underlying stress).

4.3.2 No Shift and Stress on Different Positions in Different Words

The ‘no shift/lexical’ training set of the list in (21) tests whether virtual learners acquire the trisyllabic restriction and lexical stress on different positions within

a word. It comprises six 2–5-syllable words that have stress on either the antepenult, the penult, or ultimate syllable. Stress does not shift to another syllable in the genitive case.

(21) The ‘no shift/lexical’ set:

Nominative singular		Genitive plural	
<tree+Nom.Sg.N>	[déndro]	<tree+Gen.Pl>	[déndron]
<sheep+Nom.Sg.N>	[próvato]	<sheep+Gen.Pl>	[próvaton]
<hope+Nom.Sg.F>	[elpída]	<hope+Gen.Pl>	[elpídon]
<brother+Nom.Sg.M>	[adelfós]	<brother+Gen.Pl>	[adelfón]
<almond+Nom.Sg.N>	[amígdalo]	<almond+Gen.Pl>	[amígdalon]
<headache+Nom.Sg.M>	[ponokéfalos]	<headache+Gen.Pl>	[ponokéfalon]

4.3.3 *Stress on Different Positions in Different Words and Shift to the Right*

The ‘always shift’ set of the list in (22) comprises six 2–5-syllable words that do not only differ in stress placement, but that furthermore always shift stress in the genitive, either from the antepenult to the penult, from the antepenult to the ultimate or from the penult to the ultimate. This set tests the trisyllabic restriction and lexical stress on different syllables, as does the ‘no shift/lexical’ set, but in addition, stress shifts to the right. The stress shift is not predictable: sometimes it shifts only one syllable to the right (e.g. in the case of *próvato* ~ *prováton*), and sometimes it shifts two syllables (e.g. *thálasa* ~ *thalasón*).

(22) The ‘always shift’ set:

Nominative singular		Genitive plural	
<tree+Nom.Sg.N>	[déndro]	<tree+Gen.Pl>	[dendrón]
<tomato+Nom.Sg.F>	[domáta]	<tomato+Gen.Pl>	[domatón]
<sheep+Nom.Sg.N>	[próvato]	<sheep+Gen.Pl>	[prováton]
<sea+Nom.Sg.F>	[thálasa]	<sea+Gen.Pl>	[thalasón]
<almond+Nom.Sg.N>	[amígdalo]	<almond+Gen.Pl>	[amígdalon]
<headache+Nom.Sg.M>	[ponokéfalos]	<headache+Gen.Pl>	[ponokefálon]

The data only contain unstressed nominative suffixes. The genitive plural suffix, which is the same across the genders, appears as stressed in some words and as unstressed in others, depending on the location of the stress in the nominative case and the extent of the stress shift.

4.3.4 *Shift of Stress Only Sometimes*

The ‘sometimes shift’ set of the list in (23) consists of six words (two disyllabic ones, two trisyllabic ones, one quadrisyllabic and one pentasyllabic word), of

which two show variation in the genitive plural (*próvato* and *amígdalo*, respectively).

It tests whether the virtual learners are able to acquire a pattern where certain words sometimes shift stress in the genitive case and sometimes do not. As before, the trisyllabic restriction still holds. Again, depending on the extent of the stress shift, the genitive plural suffix appears stressed in some words while it appears as unstressed in others. This set furthermore contains a word with final stress in the nominative case, *vunó*. This means in effect, that there is also variation for the nominative neuter suffix *-o*, which here appears as stressed in some words (well, one word), but not in others.

(23) The 'sometimes shift' set:

Nominative singular		Genitive plural	
<tree+Nom.Sg.N>	[déndro]	<tree+Gen.Pl>	[déndron]
<mountain+Nom.Sg.N>	[vunó]	<mountain+Gen.Pl>	[vunón]
<tomato+Nom.Sg.F>	[domáta]	<tomato+Gen.Pl>	[domatón]
<sheep+Nom.Sg.N>	[próvato]	<sheep+Gen.Pl>	[próvaton]
		<sheep+Gen.Pl>	[prováton]
<almond+Nom.Sg.N>	[amígdalo]	<almond+Gen.Pl>	[amigdálon]
		<almond+Gen.Pl>	[amígdalon]
<headache+Nom.Sg.M>	[ponokéfalos]	<headache+Gen.Pl>	[ponokefálon]

4.3.5 Shift of Stress Only in Some Words, Not in Others

The 'some shift' set in (24) consists of eight words of which four words shift stress and four do not. It tests whether the virtual learners are able to acquire a pattern where some words shift stress in the genitive case and some do not. The difference to the set in (23) is that there is no variation in the sense that a word sometimes shifts the stress in the genitive case; but rather, it either always shifts the stress or it never shifts it.

(24) The 'some shift' set:

Nominative singular		Genitive plural	
<tree+Nom.Sg.N>	[déndro]	<tree+Gen.Pl>	[déndron]
<mountain+Nom.Sg.N>	[vunó]	<mountain+Gen.Pl>	[vunón]
<tomato+Nom.Sg.F>	[domáta]	<tomato+Gen.Pl>	[domatón]
<sheep+Nom.Sg.N>	[próvato]	<sheep+Gen.Pl>	[próvaton]
<butter+Nom.Sg.N>	[vútiro]	<butter+Gen.Pl>	[vutíron]
<brother+Nom.Sg.M>	[adelfós]	<brother+Gen.Pl>	[adelfón]
<almond+Nom.Sg.N>	[amígdalo]	<almond+Gen.Pl>	[amigdálon]
<headache+Nom.Sg.M>	[ponokéfalos]	<headache+Gen.Pl>	[ponokefálon]

This set contains two words that are stressed on the ultima even in the nominative case: *vunó* and *adelfós*. They do not shift stress in the genitive case, that is, the root remains unstressed in the genitive. As a consequence, the neuter suffix *-o* and the masculine suffix *-os* appear as stressed in some words and as unstressed in others. As before, the genitive plural suffix appears as stressed in some words and as unstressed in others.

4.4 The Constraints

Three types of constraints are implemented in the learners: structural constraints evaluating the surface structures and overt forms, faithfulness constraints connecting the phonological surface structures to the underlying representations, and lexical constraints evaluating the underlying representations, comprising the phonological part of the lexicon. The structural and faithfulness constraints shared by all learners are given in (25) (constraints are defined as in Tesar and Smolensky, 2000):

(25) Constraints:

Structural constraints	Faithfulness constraints
ALL-FEET-LEFT/RIGHT (AFL/AFR)	MAXROOT, MAXAFFIX
FOOTBINARITY (FTBIN)	
IAMB	
NONFINALITY (NONFIN)	
PARSE	
TROCHEE	

AFR and AFL align the feet in words with either the right or left word edge. FTBIN requires binary feet; monosyllabic feet violate this constraint. IAMB and TROCHEE require right- or left-headed feet, respectively. NONFIN punishes words that include the final syllable in a foot, and PARSE requires syllables to be parsed in a foot. MAXROOT and MAXAFFIX require underlying stress to surface.

The number of lexical constraints is not the same for all learners and depend on which forms they are trained on in the learning phase. The lexical constraints take the form of ‘connect meaning and syntactic information to a certain phonological underlying representation’, for example, <Nom. Sg.F> |-a|, where ‘<Nom.Sg.F>’ stands for the nominative case singular feminine, and ‘|-a|’ stands for the underlying phonological representation of the suffix.

4.5 Predictions

Several predictions can be made as to how the learners of the different sets are going to perform:

- The ‘no shift/default’ learners should be able to acquire the trisyllabic window and assign stress mainly by the grammar, that is, the structural constraints should play the leading role. However, an interpretation of the forms as sometimes being lexically stressed underlyingly cannot be excluded.
- The ‘no shift/lexical’ learners should acquire the trisyllabic restriction and lexical stress on the final three syllables, depending on the word.
- The ‘always shift’ learners are expected to acquire stress on the final three syllables depending on the word, and to perform the shift in the genitive case sometimes; more often, the same stress pattern as in the nominative case should occur.
- The ‘sometimes shift’ and the ‘some shift’ learners are expected to learn lexical stress as well as the shift, but to a lesser extent than the ‘always shift’ learners, due to a less frequent occurrence of the shift.

4.6 Computer Settings

All training sets have words of a length of two to five syllables in common, as well as the trisyllabic restriction (i.e. that stress never occurs to the left of the antepenultimate syllable).

Depending on the training set the virtual learners were supplied with different sets of lexical constraints. Only lexical constraints concerning forms that were actually observed in the training sets were implemented in the virtual learners (for instance, learners of the ‘no shift/lexical’ set did not have a lexical constraint for the word ‘mountain’, <mountain> |vun-|, that the learners of the ‘sometimes’ and ‘some shift’ set had).

In both virtual comprehension and production, the learners could choose from a list of candidate quadruplets consisting of meaning, underlying representation, surface structure and overt forms. The candidates within one tableau differed with respect to the underlying stress specification (stressed/unstressed), the foot and stress structures on the surface level (iambes versus trochees, monosyllabic versus disyllabic feet, final syllable extrametricality and direction of parsing) and different stress positions on the overt level.

The virtual learners were all stochastic Harmonic-Grammar learners (Boersma and Pater, 2008) raised in the Praat programme (Boersma and Weenink, 2009).

For each training set, ten virtual learners were conceived. They all heard 200,000 items in their training phase, in a randomized order and with slightly different frequencies (because the forms were drawn randomly from the different sets and the total number of forms of each set are not always dividers of 200,000), but a rough ratio of 3:1 nominative versus genitive case forms. Plasticity, that is, the amount that constraints were shifted in case of an error, was set to 1 and did not change in the course of learning. Noise was set to 2 (enabling constraints to switch for each new evaluation of a form). The structural constraints were ranked at a 100 initially, while the faithfulness and lexical constraints were ranked at 0. This was to prevent an initial influence of the lexical constraints over the structural constraints, to simulate the creation of the lexical constraints (that were implemented but should in principle be created by the learners themselves).

4.7 Results

In all conditions, variation occurred on the underlying, the surface or the overt level. The variation occurred within all learners, not across learners.

4.7.1 *The 'No Shift/Default' Virtual Learners*

As mentioned above, the 'no shift/default' learners were trained on four different words (one disyllabic, one trisyllabic, one quadrisyllabic and one 5-syllable word) in the nominative and the genitive case (eight words in total repeated about 25,000 times each). The virtual learners were supplied with the structural and faithfulness constraints discussed in Section 4.4. Since the stress did not shift in the genitive case, seven lexical constraints were implemented: one for each root (in this case underlyingly stressed <tree> |déndr-|, <sheep> |próvat-|, <almond> |amígdal-| and <headache> |ponokéfal-|) and one for each encountered suffix (underlyingly unstressed <Nom.Sg.N> |-o|, <Nom.Sg.M.> |-os| and <Gen.Pl> |-on|). Table 5.1 shows the results for the training data (pairs of meaning and phonetic form). Learning can be called successful, since all virtual learners reproduce the data they heard to a hundred percent correct.

Variation occurred on the hidden levels that the virtual learners had to come up with themselves: the underlying representations and surface structures, respectively. The candidates that the virtual learners could choose from were quadruplets of meaning, underlying representation, surface structure with feet and the overt forms. These quadruplets differed in whether the underlying morphemes were stressed or not, and in foot structures on the surface level. For instance, for the meaning <tree>, two possible underlying representations were induced, |dendr-| and |déndr-|. ⁹In addition, all possible underlying representations were combined with surface structures with different foot structures

Table 5.1 Results of the ‘no shift/default’ condition

<i>Nominative singular</i>		%
<tree+Nom.Sg.N>	[déndro]	100
<sheep+Nom.Sg.N>	[próvato]	100
<almond+Nom.Sg.N>	[amígdalo]	100
<headache+Nom.Sg.M>	[ponokéfalos]	100
<i>Genitive plural</i>		%
<tree+Gen.Pl>	[déndron]	100
<sheep+Gen.Pl>	[próvaton]	100
<almond+Gen.Pl>	[amígdalon]	100
<headache+Gen.Pl>	[ponokéfalon]	100

and hence different positions of stress. The combined overt forms only carried the stress of the respective surface structure, that is, there is never a mismatch between the stress of a surface structure and its accompanying overt form. Table 5.6 (see appendix) shows a sample of the candidate lists for each word, namely the variation in the winning candidate quadruplets that differed only in the underlying forms (both forms are deemed equally correct because they would sound the same on the phonetic level).¹⁰ We can see that the virtual learners slightly prefer the candidates with underlyingly stressed root morphemes to the underlyingly unmarked ones. The surface structures (represented in slashes ‘/ /’) did not vary: the grammars of the virtual learners assign binary feet, even to the cost of violating final syllable extrametricality in the disyllabic words.

In short, the virtual learners were able to acquire the stress pattern and assigned stress always on the antepenultimate syllable (except for the disyllabic forms). However, they had a slight preference to assume underlyingly stressed root morphemes. At this point the question arises whether the virtual learners truly acquired the trisyllabic restriction as a phonological rule, or whether they simply assign lexical stress to every word. In anticipation of this question, the virtual learners were implemented before learning with words that matched the training forms in number of syllables, but that the virtual learners did not get to hear during the training phase. In the test phase, when prompted to produce these unheard words, it appeared that even these generalizations were correct: all unheard words were accordingly stressed, even though no lexical constraints for these words existed. They were stressed only by the weighting of the structural constraints, as can be seen in Table 5.7 (see appendix). In these

cases, the virtual learners assigned about chance probability of underlyingly stressed and unstressed representations.

The next section shows the outcome of the 'no shift/lexical' set, where the virtual learners do not only have to learn the trisyllabic restriction, but also that stress can occur on any of the last three syllables.

4.7.2 *No Shift/Lexical Stress*

In addition to the four words of the 'no shift/default' set, the 'no shift/lexical' set includes two more words: a trisyllabic word with penultimate stress and a trisyllabic word with ultimate stress. Again, stress does not shift to another syllable in the genitive form. The virtual learners of this set are supplied with 16 lexical constraints (in addition to the structural and faithfulness constraints): eight for the roots, and eight for the suffixes. The excess lexical constraints are due to the fact that the two suffixes *-os* and *-on* appear as stressed in the ultimately stressed word for 'brother', while they appear as unstressed in the other words. In addition, a feminine suffix *-a* occurs.

It can be said that the 'no shift/lexical' virtual learners were also successful in acquiring the correct stress pattern, as can be seen in Table 5.2. However, in the hidden structures more variation occurs; this time both on the surface level and the underlying level (Table 5.8, see appendix). All virtual learners chose the according underlying specification (stressed/unstressed) in the roots, as observed in the data, but alternated with respect to the specification of the genitive plural suffix, as can be expected from the training data. On the surface, feet were still invariably binary, but ranged from trochees to iambs. Extrametricality became rather unimportant with the inclusion of finally stressed words.

As can be seen from the generalizations to new forms, the trisyllabic restriction was acquired as well. However, due to the fact that lexical stress appeared on different positions, variation occurred as to where stress could fall in these words (Table 5.9, see appendix).

In sum we can state that the lexical specifications of this training condition are learned: the roots are underlyingly stressed on their respective syllable. The trisyllabic restriction is acquired in the sense that stress is never assigned pre-antepenultimately, even in new words.

However, the antepenultimate syllable is not the default position in new words. Although variation occurs even on the phonetic level, a preference for underlyingly unstressed, finally stressed forms (e.g. |stafid+a| / sta.(fi.dá) / [stafidá]) can be observed. This is surprising, because there is only one word that is stressed on the ultimate, *adelfós*.

Structurally speaking, binary feet are preferred over extrametricality and uniformity of foot type (i.e. both iambs and trochees are generated).

Table 5.2 Results of the ‘no shift/lexical’ condition

<i>Nominative singular</i>		%
<tree+Nom.Sg.N>	[déndro]	100
<sheep+Nom.Sg.N>	[próvato]	100
<hope+Nom.Sg.F>	[elpída]	100
<brother+Nom.Sg.M>	[adelfós]	100
<almond+Nom.Sg.N>	[amígdalo]	100
<headache+Nom.Sg.M>	[ponokéfalos]	100
<i>Genitive plural</i>		%
<tree+Gen.Pl>	[déndron]	100
<sheep+Gen.Pl>	[próvaton]	100
<hope+Gen.Pl>	[elpídon]	100
<brother+Gen.Pl>	[adelfón]	99.9
<almond+Gen.Pl>	[amígdalon]	100
<headache+Gen.Pl>	[ponokéfalon]	100

4.7.3 *Shifting Always*

The ‘always shift’ set tests the trisyllabic restriction and lexical stress on different syllables. The set comprises six 2–5-syllable words that do not only differ in stress placement, but that also shift stress in the genitive, either from the antepenult to the penult, from the antepenult to the ultima or from the penult to the ultima.

The virtual learners of this condition struggled to acquire the pattern, and had a considerably high error quota.¹¹ Not only do the learners have problems with learning the lexical stress of the nominative case (something the learners of the second condition had no problems with), but also some shifts in the genitive seem to be harder to acquire than others. Shifts from the penult to the ultima are acquired comparatively well, while shifts from the antepenult to the ultima or the penult are poorly acquired as shown in Table 5.3.

Even in this case, though, the trisyllabic restriction is acquired; there is no preference for a clear stress pattern in new words, but no word of the trained condition nor the untrained condition is stressed further to the left than the antepenult. These learners therefore acquire the trisyllabic restriction as well. The virtual learners hence have more difficulty in acquiring the ‘always shifting’ pattern.

Table 5.3 Results of the ‘always shift’ condition

<i>Nominative singular</i>		%
<tree+Nom.Sg.N>	[déndro]	89.0
<tomato+Nom.Sg.F>	[domáta]	82.2
<sheep+Nom.Sg.N>	[próvato]	72.5
<sea+Nom.Sg.F>	[thálasa]	68.4
<almond+Nom.Sg.N>	[amígдалo]	80.0
<headache+Nom.Sg.M>	[ponokéfalos]	68.5
<i>Genitive plural</i>		%
<tree+Gen.Pl>	[dendrón]	93.6
<tomato+Gen.Pl>	[domatón]	45.0
<sheep+Gen.Pl>	[provátón]	12.8
<sea+Gen.Pl>	[thalasón]	2.6
<almond+Gen.Pl>	[amigdálon]	16.1
<headache+Gen.Pl>	[ponokefálon]	28.8

4.7.4 *Shifting Sometimes*

The ‘sometimes shift’ set (Table 5.4) tests whether the virtual learners are able to acquire a pattern where certain words sometimes shift stress in the genitive case and sometimes not. The set consists of six words, of which two show variation in the genitive plural. This pattern is acquired better than the ‘always shifting’ condition, but compared to the nominative case, the genitive case is considerably badly acquired.

In the ‘sometimes shift’ condition, the non-shifting forms were better acquired than the shifting forms. This looks like it could not be a frequency effect since both genitive variants were presented equally often, but the non-shifting stress genitive cases add up to the non-shifting stress forms of the nominative case. This result is in contrast to the small sample count of the HNC, where shifting genitive forms outnumbered the non-shifting genitive forms. Even in the untrained forms though, the trisyllabic restriction is obeyed.

4.7.5 *Some Shift, Some Don’t*

As one can see in Table 5.5, the learners of the ‘some shift’ set behave similar to the ‘sometimes shift’ learners, in that they acquire the nominative case forms better than the genitive case forms. Words that shift their stress in the genitive case are furthermore produced less often correct than the nominative forms of not-shifting genitive forms.

Table 5.4 Results of the ‘sometimes shift’ condition

<i>Nominative singular</i>		%	
<tree+Nom.Sg.N>	[déndro]	100	
<mountain+Nom.Sg.N>	[vunó]	100	
<tomato+Nom.Sg.F>	[domáta]	93.5	
<sheep+Nom.Sg.N>	[próvato]	85	
<almond+Nom.Sg.N>	[amígdalo]	80.8	
<headache+Nom.Sg.M>	[ponokéfalos]	76.2	
<i>Genitive plural</i>		%	
<tree+Gen.Pl>	[déndron]	100	
<mountain+Gen.Pl>	[vunón]	100	
<tomato+Gen.Pl>	[domatón]	50.9	
<sheep+Gen.Pl>	[próvaton]	71.2	[provátón] 6.7
<almond+Gen.Pl>	[amígdalon]	64.1	[amígdálon] 12
<headache+Gen.Pl>	[ponokefálon]	23.9	

Table 5.5 Results of the ‘some shift’ condition

<i>Nominative singular</i>		%	
<tree+Nom.Sg.N>	[déndro]	100	
<mountain+Nom.Sg.N>	[vunó]	100	
<tomato+Nom.Sg.F>	[domáta]	99.9	
<sheep+Nom.Sg.N>	[próvato]	100	
<butter+Nom.Sg.N>	[vútiro]	66.1	
<brother+Nom.Sg.M>	[adelfós]	99.3	
<almond+Nom.Sg.N>	[amígdalo]	80.2	
<headache+Nom.Sg.M>	[ponokéfalos]	72.4	
<i>Genitive plural</i>		%	
<tree+Gen.Pl>	[déndron]	100	
<mountain+Gen.Pl>	[vunón]	100	
<tomato+Gen.Pl>	[domatón]	80.3	
<sheep+Gen.Pl>	[próvaton]	99.9	
<butter+Gen.Pl>	[vutíron]	32.6	
<brother+Gen.Pl>	[adelfón]	80.4	
<almond+Gen.Pl>	[amígdálon]	16.8	
<headache+Gen.Pl>	[ponokefálon]	27.8	

New words were produced with mainly stress on the penult in the nominative case and on the ultima in the genitive case, with the exception of words ending in *-o* and *-os*: since they occurred as stressed in some nominative case forms, they tilted the results to final stress in the nominative case in words with these suffixes. These learners as well acquired the trisyllabic restriction (a full list is omitted for space reasons).

4.8 Resulting Weights

In all conditions, AFR ended up with the highest weight (way above AFL), and the genitive plural unstressed allomorph |-on| with one of the lowest. FRBIN and PARSE never moved, indicating that they were never violated by a winning candidate (hence no error detection that could have elicited a re-weighting). This could have been due to the fact that they started out with higher initial ranking than the faithfulness and lexical constraints. TROCHEE always ended up higher than NONFINAL and IAMB, and MAXR always ended up higher than MAXA. The weights of the structural constraints caused binary feet in all surface forms across learners and conditions, to the cost of syllable extrametricality and uniformity of rhythm (the learners both exhibited trochees and iambs in their surface forms). Together, the weights of the structural constraints ensured the preservation of the trisyllabic window across all conditions.

In the 'no-shift/default' condition, the lexical constraints hardly moved, as can be expected, since in this condition, stress was supposed to be assigned phonologically and not lexically (Table 5.10, see appendix).

In the 'no-shift/lexical' condition, the lexical constraints pertaining to stressed allomorphs were weighted higher than most of the lexical constraints pertaining to unstressed allomorphs, reflecting the fact that in this condition, stress should have been underlyingly marked (Table 5.11, see appendix).

In the 'always shift' condition, the lexical constraints ended up with weights still close to 0, their starting point. This is due to the fact that the training data contained conflicting items, and therefore led to the nominative case being less well acquired than in the 'no shift/lexical' condition (which differed only in the genitive case from the 'always shift' condition). The genitive case shift was even more poorly acquired probably because of its lesser frequency in the training data. The shift across two syllables in *thálasa* ~ *thalasón* is especially poorly acquired; this could also be a frequency effect (Table 5.12, see appendix).

In the 'sometimes shift' condition, the main difference to the 'always shift' condition is that the lexical constraints for the nominative suffixes have moved, resulting in an even worse acquisition of the shift (Table 5.13, see appendix).

In the 'some shift' condition, the weight of MAXA is much lower than in the other conditions, in relation to the other constraints. This is connected to the

lexical constraints for the unstressed affix-allomorphs, which end up ranked, in opposition to the other conditions, where these lexical constraints often remained unranked (Table 5.14, see appendix).

In general, lexical constraints pertaining to unstressed allomorphs often did not move in the course of learning.

5. Discussion and Conclusion

All virtual learners were able to acquire the trisyllabic restriction even for words that they had not been trained on. However, neither stress on the antepenultimate (as proposed by, for example, Philippaki-Warbuton, 1970) nor the penultimate syllable (as proposed by Protopapas et al., 2006) were the default pattern for untrained forms; rather, the final syllable was a preferred position. Apart from that, the stress patterns were acquired with different degrees of success. The no shift/default and the no shift/lexical pattern were acquired correctly; the child generation did not differ in their overt forms from the parent generation. In the former pattern, the structural constraints are responsible for stress assignment, in the latter, mainly the lexicon.

The patterns including a shift in the genitive case were less successfully acquired. While stress on different positions within words was acquired less well than in the no-shift conditions, it was better acquired than the shifts. Among the shifts, the ones from the penult to the ultima were acquired better than shifts from the antepenult to the ultima or the penult. Moreover, the nominative case forms were more poorly acquired in the always-shift condition than in the two hybrid patterns, where the nominative case forms were almost always produced correctly. While an in-depth analysis of the learning paths could shed some light on this issue, it is unlikely that frequency was the cause for this effect. The virtual learners were exposed to nominative versus genitive forms in a 3:1 ratio, however, the genitive plural suffix occurred much more often than any of the roots or other suffixes simply because it was the only one for the genitive case, while for the nominative case different suffixes occurred depending on the gender of the form. Raising the occurrence of the genitive case (let's say, to a 1:1 ratio) would therefore probably not improve the learnability of the shift, but only worsen the learnability of the nominative case.

Transferring the results of the computational simulations to spoken Greek, this would mean that the variation in the spoken language is due to an irreconcilability of grammatical and lexical demands of the different patterns, resulting from the Ancient Greek incomplete influence. An empirical study of the variation in real Modern Greek (i.e. whether speakers have problems in producing the genitive shift in general, or whether they only have problems with some shifts) could shed more light on how far the proposed model is correct.

6. Acknowledgements

I thank Paul Boersma and Joe Pater for helpful discussions on some early ideas of this chapter; any errors and misconceptions are nonetheless my very own. This work is furthermore owed to Rubicon grant No. 446–07-030, funded by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

7. Appendix

Table 5.6 ‘No shift/default’-results with hidden structure: Underlying variation

<i>Output</i>				<i>%</i>
<tree+Nom.Sg.N>	dendr+o	/ (dén.dro) /	[déndro]	44.4
	déndr+o			55.6
<tree+Gen.Pl>	dendr+on	/ (dén.dron) /	[déndron]	44.2
	déndr+on			55.8
<sheep+Nom.Sg.N>	provat+o	/ (pró.va).to /	[próvato]	41.5
	próvat+o			58.5
<sheep+Gen.Pl>	provat+on	/ (pró.va).ton /	[próvaton]	41.3
	próvat+on			58.7
<almond+Nom.Sg.N>	amígdal+o	/ a.(míg.da).lo /	[amígdalo]	50.2
	amígdal+o			49.8
<almond+Gen.Pl>	amígdal+on	/ a.(míg.da).lon /	[amígdalon]	50.2
	amígdal+on			49.8
<headache+Nom.Sg.M>	ponokefal+os	/ po.no.(ké.fa).los /	[ponokéfalos]	25.8
	ponokéfal+os			74.2
<headache+Gen.Pl>	ponokefal+on	/ po.no.(ké.fa).lon /	[ponokéfalon]	25.6
	ponokéfal+on			74.4

Table 5.7 'No shift/default'-results of generalizations to new words: More underlying variation

<i>Output</i>				<i>%</i>
<bath+Nom.Sg.N>	banj+o	/(bán.jo)/	[bánjo]	49.9
	bánj+o	/(bán.jo)/		50.1
<bath+Gen.Pl>	banj+on	/(bán.jon)/	[bánjon]	49.9
	bánj+on			50.1
<butter+Nom.Sg.N>	vútir+o	/(vú.ti).ro/	[vútiro]	50.2
	vútir+o			49.8
<butter+Gen.Pl>	vútir+on	/(vú.ti).ron/	[vútiron]	50.2
	vútir+on			49.8
<smile+Nom.Sg.N>	xamojel+o	/xa.(mó.je).lo/	[xamójelo]	49.9
	xamójel+o			50.1
<smile+Gen.Pl>	xamojel+on	/xa.(mó.je).lon/	[xamójelon]	50
	xamójel+on			50
<railway+Nom.Sg.M>	sidirodrom+os	/si.di.(ró.dro).mos/	[sidiródromos]	50.2
	sidiródrom+os			49.8
<railway+Gen.Pl>	sidirodrom+on	/si.di.(ró.dro).mon/	[sidiródromon]	49.9
	sidiródrom+on			50.1

Table 5.8 'No shift/lexical'-results: Underlying and surface variation

<i>Output</i>				%
<tree+Nom.Sg.N>	déndr+o	/(dén.dro)/	[déndro]	100
<tree+Gen.Pl>	déndr+on déndr+ón	/(dén.dron)/	[déndron]	27.8 72.2
<sheep+Nom.Sg.N>	próvat+o	/(pró.va).to/	[próvato]	100
<sheep+Gen.Pl>	próvat+on próvat+ón	/(pró.va).ton/	[próvaton]	27.9 72.1
<hope+Nom.Sg.F>	elpíd+a	/(el.pí).da/ /el.(pí.da)/	[elpída]	29.9 70.1
<hope+Gen.Pl>	elpíd+on elpíd+ón	/(el.pí).don/ /el.(pí.don)/ /(el.pí).don/ /el.(pí.don)/	[elpídon]	6.5 21.5 23.3 48.7
<brother+Nom.Sg.M>	adelf+ós	/a.(del.fós)/	[adelfós]	100
<brother+Gen.Pl>	adelf+ón	/a.(del.fón)/	[adelfón]	99.9
<almond+Nom.Sg.N>	amígdal+o	/a.(míg.da).lo/	[amígdalo]	100
<almond+Gen.Pl>	amígdal+on amígdal+ón	/a.(míg.da).lon/	[amígdalon]	27.9 72.1
<headache+Nom.Sg.M>	ponokéfal+os ponokéfal+ós	/po.no.(ké.fa).los/	[ponokéfalos]	5.9 94.1
<headache+Gen.Pl>	ponokéfal+on	/po.no.(ké.fa).lon/	[ponokéfalon]	100

Table 5.9 ‘No shift/lexical’-results of generalizations to new words:
More underlying and surface variation

<i>Output</i>				%
<bath+Nom.Sg.N>	banj+o	/b(án.jo)/	[bánjo]	17.1
		/(ban.jó)/	[banjó]	58.6
	bánj+o	/b(án.jo)/	[bánjo]	17.2
		/(ban.jó)/	[banjó]	7.1
<bath+Gen.Pl>	banj+ón	/(ban.jón)/	[banjón]	100
<butter+Nom.Sg.N>	vutir+o	/v(ú.ti).ro/	[vútiro]	2.4
		/(vu.tí).ro/	[vutíro]	9.8
		/vu.(tí.ro)/	[vutíro]	29.7
		/vu.(ti.ró)/	[vutiró]	49.3
	vútir+o	/v(ú.ti).ro/	[vútiro]	2.4
		/vu.(ti.ró)/	[vutiró]	6.4
<butter+Gen.Pl>	vutir+ón	/vu.(ti.rón)/	[vutirón]	99.9
<raisin+Nom.Sg.F>	stafid+a	/(stá.fi).da/	[stáfida]	4.3
		/(sta.fí).da/	[stáfida]	4.1
		/sta.(fí.da)/	[stáfida]	12.1
		/sta.(fi.dá)/	[stafidá]	27.2
	stafíd+a	/(sta.fí).da/	[stáfida]	4
		/sta.(fí.da)/	[stáfida]	12.0
		/sta.(fi.dá)/	[stafidá]	36.3
<raisin+Gen.Pl>	stafid+ón	/sta.(fi.dón)/	[stafidón]	99.9
<sky+Nom.Sg.M>	uran+ós	/u.(ra.nós)/	[uranós]	100
<sky+Gen.Pl>	uran+ón	/u.(ra.nón)/	[uranón]	99.9
<smile+Nom.Sg.N>	xamojel+o	/xa.(mó.je).lo/	[xamójelo]	2.4
		/xa.(mo.jé).lo/	[xamójelo]	9.8
		/xa.mo.(jé.lo)/	[xamójelo]	29.8
		/xa.mo.(je.ló)/	[xamojeló]	49.3
	xamójel+o	/xa.(mó.je).lo/	[xamójelo]	2.4
		/xa.mo.(je.ló)/	[xamojeló]	6.3
<smile+Gen.Pl>	xamojel+ón	/xa.mo.(je.lón)/	[xamojelón]	99.9
<railway+Nom.Sg.M>	sidirodrom+ós	/si.di.ro.(dro.mós)/	[sidirodromós]	100
<railway+Gen.Pl>	sidirodrom+ón	/si.di.ro.(dro.món)/	[sidirodromón]	99.9

Table 5.10 Weightings for the ‘no shift/default’ condition

<i>Constraints</i>	<i>weights</i>
AFR	107
TROCH	106.8
{FTBIN, NONFINAL; PARSE}	100
IAMB	93.2
AFL	93
MAXR	3.8
<headache> ponokéfal-	1.5
<sheep> próvat-	0.5
<tree> déndr-	0.3
{MAXA; <Nom.Sg> -a , -o , -os ; <Gen.Pl> -on }	0
<almond> amígdal-	0

Table 5.11 Weightings for the ‘no shift/lexical’ condition

<i>Constraints</i>	<i>weights</i>
AFR	134.7
IAMB	100.6
{FTBIN; PARSE}	100
TROCH	99.4
NONFINAL	65.6
AFL	65.3
MAXR	39.5
<headache> ponokéfal-	39.5
<almond> amígdal-	36.7
<sheep> próvat-	36.4
<tree> déndr-	27.5
<hope> elpíd-	26.3
MAXA	12.5
<Nom.Sg.M> -ós	9.2
<Gen.Pl> -ón	7.5
<Nom.Sg.N> -o	2
{<Nom.Sg.F> -a , -as ; <mountain> vun- ; <tomato> domát- ; <brother> adelf- }	0
<Nom.Sg.N> -ó	-2
<Gen.Pl> -on	-7.5
<Nom.Sg.M> -os	-9.2

Table 5.12 Weightings for the ‘always shift’ condition

<i>Constraints</i>	<i>weights</i>
AFR	132.1
TROCH	101.4
{FTBIN, PARSE}	100
IAMB	98.6
NONFINAL	68.9
AFL	67.9
MAXR	17.3
MAXA	14.5
<Gen.Pl> -ón	11.8
<sheep> próvat-	6.3
<almond> amígdal-	5.8
<sea> thalas-	5.5
<tomato> domát-	5.3
<almond> amigdál-	5.2
<sheep> provát-	4.8
{<Nom.Sg> -o ; -as ; -os ; -a }	0
<tree> déndr-	0 ¹³
<tomato> domat-	-5.3
<sea> thálas-	-5.5
<headache> ponokéfal-	-7.3
<headache> ponokefál-	-7.7
<Gen.Pl> -on	-11.8

Table 5.13 Weightings for the ‘sometimes shift’ condition

<i>Constraints</i>	<i>weights</i>
AFR	127.8
IAMB	104.4
{FTBIN; PARSE}	100
TROCH	95.6
NONFIN	73.2
AFL	72.2
<tree> déndr-	42.4
MAXR	42.3
MAXA	18.1
<almond> amígdal-	15.4
<Gen.Pl> -ón	14
<sheep> próvat-	13.8
<headache> ponokéfál-	10.4
<mountain> vun-	5.4
<tomato> domat-	4.6
<Nom.Sg.F> -a	4.2
<Nom.Sg.M> -os	2.9
<Nom.Sg.N> -o	1.3
{<Nom.Sg.M> -as ; <hope> elpíd- ; <brother> adelf- }	0
<Nom.Sg.N> -ó	-1.3
<almond> amigdál-	-2.2
<Nom.Sg.M> -ós	-2.9
<tomato> domát-	-4.6
<sheep> provát-	-5
<headache> ponokefál-	-7.5
<Gen.Pl> -on	-14

Table 5.14 Weightings for the ‘some shift’ condition

<i>Constraints</i>	<i>weights</i>
AFR	129.2
IAMB	105.8
{FTBIN, PARSE}	100
TROCH	94.2
NONFINAL	71.2
AFL	70.8
MAXR	47.9
<tree> déndr-	47.2
<headache> ponokéfal-	42.5
<sheep> próvat-	40.1
<butter> vútír-	32
<almond> amígdál-	29.2
<headache> ponokéfal-	27.6
MAXA	26.3
<butter> vutír-	17.1
<Nom.Sg.M> -ós	15.5
<almond> amígdál-	11.2
<Gen.Pl> -ón	11
<Nom.Sg.F> -a	5.5
<tomato> domát-	4.7
<Nom.Sg.N> -o	3.2
<Nom.Sg.M> -as ; <mountain> vun- ; <hope> elpíd- ; <brother> adelf-	0
<Nom.Sg.N> -ó	-3.2
<tomato> domát-	-4.7
<Gen.Pl> -on	-11
<Nom.Sg.M> -os	-15.5

8. Notes

1. The expression 'lexical stress' is sometimes used as referring to word stress in contrast to sentence stress.
2. This seems to hold more for the right edge than the left edge of words; post-postinitial stress is so far missing from the typology. See Hyman (1977a) for an indication of frequent and infrequent edge-oriented patterns.
3. See Dogil et al. (1999: 852) for a phonological rather than lexical account of Russian stress.
4. Apparently, loanwords can violate the trisyllabic window, for example, *kámeraman* 'cameraman' (Revithiadou, 1999: 95).
5. See Protopapas et al. (2006) and references therein for suggesting that the default is actually stress on the penultimate syllable, based on the frequency of this pattern.
6. The stories by Allen (1689/1999) and Devine and Stephens (1994) go that the length distinction was lost because the duration of the long vowels was re-interpreted as a stress cue, with a subsequent shortening of all unaccented vowels.
7. These assumptions are certainly not trivial; the learner already needs to know what segments are, needs to be able to hear stress, needs to know the meaning, needs to know that there are two morphemes in the word etc. However, I leave the modelling of the acquisition of this knowledge is, at this point, for future research.
8. 'Nom.' stands for nominative case; 'Gen.' for genitive; 'Sg.' for singular; 'Pl.' for plural; and 'M', 'F' and 'N' for masculine, feminine and neuter gender, respectively.
9. Although not observed in this training set, unstressed underlying forms were always included as possible candidate parts as the most unmarked representations.
10. The complete list of all assumed candidates is omitted for reasons of space.
11. It is possible that with even more training, the learners would become a little bit better, but considering the amount of training data and the rather big learning steps, this seems not likely.

6 Derivations

Patrik Bye

Chapter Overview

1. Derivations and Levels of Representation	135
2. Derivations in Phonological Theory	140
3. Alternatives to Opaque Interaction	165
4. Conclusions	170
5. Acknowledgements	171
6. Notes	171

1. Derivations and Levels of Representation

Two intimately related concepts in linguistics are **level** and **derivation**. Generative phonology recognizes a minimum of two levels of representation, generally known as the **underlying** and **surface** levels. The derivation is the mapping from underlying to surface representation. So far there is nothing uniquely generative about this conception. Levels and mappings between levels are known from structuralism, so it is important to understand how structuralist and generative notions of language differ in order to get an adequate grasp of both of these concepts.

At its broadest, the *level* in structuralist analysis refers to a way of describing a linguistic expression from a particular point of view. A sentence like *the Celts wore kilts*, for example, can be described from a syntactic perspective [[the Celts] [wore [kilts]]], morphological {determiner N+plural V+past N+plural}, and phonetic [ðə 'kɛlts wə: 'kɪlts]. Each of these descriptions is autonomous, with its own vocabulary and principles of well-formedness. In describing the sound pattern of a language, structuralism minimally invoked two levels of description, termed **phonemic** and **phonetic** (a.k.a. **allophonic**). The phonetic level

supplies a more or less detailed transcription of an utterance understood as an individual speech event at a particular time and place, or an equivalence class of such speech events. A narrow transcription of our sentence might be [ðə 'k^hɛt^s wɔ: k^hɪt^s]. The phonemic level abstracts away from non-contrastive variation in the speech signal such as post-aspiration, the velarization of laterals, or precise shades of vowel quality. A phonemic transcription of the same sentence might therefore be /ðə kelts wɔ: kilts/. The end result of phonemic analysis is a set of autonomous phonemes along with a statement of their allophones and the contexts in which they appear. Beyond this, it is possible to distinguish three different philosophical positions on the phoneme that we can designate by the terms **psychological realism** or **mentalism**, **physical realism** and **nominalism**. Linguists such as Edward Sapir (1925), Nikolai Trubetzkoy and Baudouin de Courtenay (Anderson, 1985: 290) believed that the phoneme could be understood as the mental image of a speech sound. The second position, associated with Leonard Bloomfield (1933) and Daniel Jones (1962), was that the phoneme was the acoustically invariant part of its allophones. The third position, first articulated by Twaddell (1935), was that the phoneme didn't refer to an objective reality at all; phonemes are merely analytical conveniences posited by the linguist so as to be able to describe the contrasts of a particular language. It was this view that came to serve as the basis for American Structuralism, which was the immediate context for the development of generative phonology.

The shift away from these structuralist conceptions began in the 1920s as distinctive feature theory began to take shape with the collaboration of Nikolai Trubetzkoy and Roman Jakobson in Prague (Joseph, 2000). Their work on neutralization ended up breaking with Saussure's purely relational conception of the phonological system as a system of unanalysable phonemes (Saussure, 1995 [1916]). The classic example, from German, is the neutralization of the phonemic distinction between /t/ and /d/ to [t] in word-final position, as illustrated by alternations such as *Rad* [ra:t] 'wheel' versus the corresponding dative form *Rade* [ra:də] (cf. *Rat* [ra:t] 'counsel', with dative *Rate* [ra:tə]). Because they exclusively alternated, /t/ and /d/ were shown to be clearly more closely related to each other, or 'correlated', than with other phonemes. Pairs of correlating phonemes were grouped into **archiphonemes** each with a **marked** and an **unmarked** member distinguished by a **principum divisionis**, in this case voicing. Neutralization also had a particular direction: it always entailed the loss of the principum divisionis so the result of the neutralization was always the unmarked member of the pair. This complex of ideas represents the beginning of reasoning phonologically in terms of natural classes, distinctive features and markedness, as well as dynamic rules.

In the years leading up to the inception of the generative paradigm in phonology, structuralist phonological analysis recognized a further **morphophonemic**

level (Swadesh and Voegelin, 1939). Just as the phonemic level abstracted away from variation due to phonetic context, the morphophonemic level was conceived to abstract away from alternations conditioned by the phonological or morphological context to permit a unique representation for each morpheme. For example, in order to handle the difference between non-alternating singular/plural pairs such as *hive* ~ *hives*, *fife* ~ *fifes* and alternating pairs like *wife* ~ *wives*, a morphophoneme /f^M/ was posited to cover cases where the phonemes /f/ and /v/ alternated. This allowed the formulation of statements specifying how morphophonemes mapped onto phonemes in particular morphophonological contexts. While recognizing morphophonemes as a convenient descriptive device, Hockett (1961) noted that the morphophonemic level was not itself autonomous, but was ‘produced [. . .] by conflicts between the independent workings of grammar and of phonology’ (p. 50). According to the structuralist conception, it was the phonemic level of description that was autonomous. With the advent of generative phonology, however, this understanding ended up being turned on its head. Beginning with Chomsky (1951), it was the morphophonemic level, or something approximating it (now generally known as the ‘underlying level’), that came to be seen as autonomous, and the **autonomous phonemic level** of structuralism was abandoned. This was part and parcel of a fundamental shift in the understanding of what linguistic symbols represented. In contrast to the structuralist levels of analysis, the underlying and surface levels in generative phonology are both made up of the same discrete symbolic units (distinctive features, segments and so on). These units purport to be psychologically real and have **intrinsic phonetic content**, as opposed to abstract phonetically uninterpretable morphophonemes.¹ The underlying and surface levels occupy unique positions in the grammar as a whole. The underlying form is what is encoded in the **lexicon**. The surface level, on the other hand, serves as an interface to the **articulatory** and **auditory systems**.

The reason for the paradigm shift can be traced to two crucial demonstrations that positing an autonomous phonemic level resulted in missing **linguistically significant generalizations** in some cases. Halle (1959) showed first that an autonomous phonemic level entailed duplicating generalizations when the effect of a rule was neutralizing in some cases and non-neutralizing in others. Russian obstruents generally contrast in voicing, as shown by minimal pairs such as /píli/ ‘dust-GEN.SG’ versus /bíli/ ‘be-PAST.PL’, and /got/ ‘year’ versus /kot/ ‘tomcat’. An obstruent assimilates in voicing to a following obstruent, for example, /k domu/ → [gdomu] ‘towards the house’, /spas bi/ → [spazbi] ‘would have saved’, /ot’ěts bil/ → [at’édzbi] ‘father was’, /dotʃ bilá/ → [dodʒbilá] ‘daughter was’. Because Russian has a voicing contrast in obstruents, obstruent voicing assimilation is generally neutralizing, but in the case of /ts/ and /tʃ/, which lack voiced counterparts, it is not. On a structuralist analysis, the neutralizing part

of the rule would have to be stated as a mapping from the morphophonemic to the phonemic level, while the non-neutralizing part would have to be stated as a mapping from the phonemic to the allophonic level.

A few years later, Chomsky (1964) showed that phonological rules are not simply directed mappings from an underlying to a surface representation. Phonological rules interact and must therefore be linearly ordered, and failing to recognize this can lead to positing false contrasts. In Canadian English (Joos, 1942), /āy/ is raised to [ʌy] before voiceless stops, and /t d/ are realized as a flap intervocalically. Because /t/ is an obstruent stop, the contexts of these rules overlap, which makes it necessary to assign priority of application to one rule or the other. As a word like *writer* [rʌyɾə] shows, raising must apply before flapping. Applying the rules in the opposite order would erroneously predict [rāyɾə], which, in the dialect in question, corresponds to the pronunciation of a different word, *rider*. The crucial ordering of these two rules thus results in **surface minimal pairs** which, if used as the sole basis for determining contrast, can overdetermine the underlying inventory. Even though the diphthongs [āy] and [ʌy] contrast exclusively before [r], the structuralist procedure for discovering phonemes, which, precisely, proceeds on the basis of (surface) minimal pairs like [rʌyɾə] versus [rāyɾə], would force us to posit two phonemes, /āy/ and /ʌy/. What the taxonomic account fails to explain is how it comes about that [ʌy] is distinct from [āy] preceding exactly those flaps that alternate with a voiceless obstruent /t/.

Furthermore, as conceived by generative phonology, the processes involved in deriving the output are not restricted to adding information, such as velarization of an underlying /l/, spreading tone to tonally unspecified vowels, building syllables over unparsed segments, and so on. Processes may also bring about destructive alterations of the underlying representation, for example, by deleting segments or changing feature specifications. In a theory admitting only binary features, all neutralization rules would be destructive. This destructive capacity is a crucial difference between derivational and non-derivational theories.

These developments led to a shift of interest away from levels and their properties to rules and the way they interact in derivations. Goldsmith (1993: 25), writing of developments in the late 1960s goes so far as to say that 'the notion of derivation changed, in many linguists' perception from being an account of the fundamental problem of levels in linguistics to being the essence itself of a linguistic analysis'.² In early generative phonology, the underlying and surface representations exert no influence on the outcome of the derivation, except in the sense that the underlying representation provides the input to the first rule. As the derivation progresses, the system retains no memory of earlier states, nor is it guided by representations of desirable future states, or functional goals. What ultimately replaced the level-based view was 'a dynamic model of linguistic analysis, in terms of which one representation is successively changed into another in a sequence that in its entirety is the account of

the expression in question'. Nevertheless, throughout the history of the field's development, phonologists have repeatedly come back to the role and properties of levels in their theories. This interest manifests itself in a variety of different ways that always reflect the formal technology of the time. We shall see several examples of this in the following pages.

By way of rounding off this section, let us take a closer look at the question which phonological theories should be considered derivational. Derivational theories make use of at least two consubstantial levels (underlying and surface) and potentially destructive processes that change the one into the other. This definition rules out **monostratal** theories, which only recognize a single level of representation. Declarative Phonology (Scobbie, 1991a, 1993; Scobbie et al., 1996) is restricted to stating generalizations over surface forms, and for this reason neutralization poses a considerable challenge. Constraints in Declarative Phonology may be viewed as either structure-checking filters, which rule representations as either grammatical or ungrammatical, or as purely structure-building operations that combine information with information already contained in the input. However, any grammar consisting solely of feature-filling processes has a monostratal interpretation, and vice versa. The original version of Optimality Theory (OT) (Prince and Smolensky, 2004 [1993]; McCarthy and Prince, 1993b) may also be understood as monostratal, since it assumes the Principle of Containment, according to which nothing is literally deleted or changed in the mapping from input to candidate output. Each output candidate simply offers an alternative parse of the input; incorporating earlier ideas (see esp. Itô, 1989), inaudible material simply remains syllabically unparsed rather than being deleted outright, while epenthesis is understood as the insertion of segmentally unfilled nuclei. OT with Correspondence Theory (McCarthy and Prince, 1995, 1999), however, since it countenances feature-changing processes, is derivational. The main difference between rule-based phonology and OT with Correspondence is whether linguistically significant generalizations reside in constraints on representations or in the rules that map one representation onto another.

Also plainly non-derivational are theories where the levels are made up of different kinds of things. An example is Firthian prosodic analysis (Firth, 1948; Ogden, 1993), where the discrete units of symbolic analysis, known as 'prosodies', are mapped onto a continuous parametric phonetic representation.

Derivational is often popularly understood as synonymous with **serial**. However, as we shall see examples below, processes may also apply simultaneously or in parallel in a derivational theory. Finally, there is the perennial question to what extent linguistically significant generalizations may be said to inhere in levels of representation or the mapping between levels. Early rule-based theory and Reiss' (2008) rule-only approach lays this burden entirely on the derivation. In monostratal theories and in Classical OT, on the other hand, generalizations

reside exclusively in constraints on the surface representation. Intermediate positions are possible. Level-based theories such as Two-Level Morphology (Koskeniemi, 1983; Karttunen, 1993), Harmonic Phonology (Goldsmith, 1993) and Cognitive Phonology (Lakoff, 1993) posit constraints both on privileged levels of representations (the number of which varies depending on the proposal) and on mappings between levels. These are construed as essentially static correspondences, lacking the interactions that characterize the derivations of rule-based theory. Although there are important points of similarity between OT and level-based theories (OT's output-orientedness), OT with Correspondence Theory works with processes. Late rule-based theory includes constraints on input and output, and later developments in OT may make additional appeal to derivational notions.

In the next two sections we will look at the interactions constitutive of derivations, and how phonologists have tried to argue for limiting them, often at the cost of descriptive adequacy. Section 2 examines the interaction of rules and processes and gives examples of their effects on surface representation. This section also discusses theories of rule and process ordering. Section 3 considers phonological and non-phonological alternatives for dealing with the empirical residue in theories excluding certain types of interaction. Section 4 presents the main conclusions.

2. Derivations in Phonological Theory

In this section, we look at theories of rule and process ordering. Section 2.1 introduces **extrinsic ordering**, where rules apply in a strict linear sequence determined by the grammar. Extrinsic ordering allows rules to interact in potentially complex ways. Section 2.2 discusses the implications of rule interaction for the abstractness of underlying representations. In Section 2.3 we examine the issue whether certain types of interaction are more natural than others. Section 2.5 looks at theories that eschew the extrinsic ordering of rules and processes, including OT in its parallel and harmonic serial implementations. The last two sections look at **intrinsic ordering**. Section 2.6 looks at within-cycle intrinsic ordering (the Elsewhere Principle), and Section 2.7 addresses the way in which the construction of morphological and syntactic constituents may interact with the phonology.

2.1 Extrinsic Rule Ordering and Transient Rules

In SPE, derivations are viewed as a series of ordered **rewrite rules** of the form $A \rightarrow B/C_D$, 'A is rewritten as B in the environment C_D (between C and D)'

(Chomsky and Halle, 1968; Bromberger and Halle, 1989). Rules specify a process or **structural change** ($A \rightarrow B$) to a **focus** (here, A) in a given **environment**³ (here, C_D). Some rules are **context-free**, specifying only a focus and a structural change, for example, $A \rightarrow B$. The union of focus and any environment is known as the **structural description**. The structural description for our hypothetical rule is thus $/CAD/$. Since rules apply in a strict linear order, each rule only has one chance to apply (or, in a model that recognizes cycles, one chance in a given cycle). To use a term from Chafe (1968), each rule is **transient**.⁴ During the 1970s and 1980s, this rule-based approach was elaborated with additional formal devices, including inventories, Morpheme Structure Conditions (MSCs), filters and constraints, as well as highly articulated theories of non-linear representations (Autosegmental Phonology, Feature Geometry and Metrical Phonology). A representative example of how rule-based phonology had come to evolve by the beginning of the 1990s is Kenstowicz (1994).

Given two **mutually affecting** rules, \mathcal{P} and \mathcal{Q} , either \mathcal{P} creates strings that satisfy the structural description of \mathcal{Q} , or \mathcal{P} destroys strings that satisfy the structural description of \mathcal{Q} . We can now define the following **two-rule interactions** in (1).⁵ The symbol $>$ is to be read 'precedes'.

- (1) a. If \mathcal{P} creates strings that satisfy the structural description of \mathcal{Q} , and both \mathcal{P} and \mathcal{Q} apply, then $\mathcal{P} > \mathcal{Q}$ and we say \mathcal{P} **feeds** \mathcal{Q} .
- b. If \mathcal{P} destroys strings that satisfy the structural description of \mathcal{Q} , and \mathcal{P} applies and \mathcal{Q} does not apply, then $\mathcal{P} > \mathcal{Q}$ and we say \mathcal{P} **bleeds** \mathcal{Q} .
- c. If \mathcal{P} creates strings that satisfy the structural description of \mathcal{Q} , and \mathcal{P} applies and \mathcal{Q} does not apply, then $\mathcal{Q} > \mathcal{P}$ and we say \mathcal{P} **counterfeeds** (fails to feed) \mathcal{Q} .
- d. If \mathcal{P} destroys strings that satisfy the structural description of \mathcal{Q} , and both \mathcal{P} and \mathcal{Q} apply, then $\mathcal{Q} > \mathcal{P}$ and we say \mathcal{P} **counterbleeds** (fails to bleed) \mathcal{Q} .

In extrinsic ordering theories, the derivation may create **intermediate representations**. Intermediate representations occur in feeding and counterbleeding interactions. \mathcal{P} creates an intermediate representation to which \mathcal{Q} then applies. Some phonologists may speak of intermediate representations as 'levels', but unlike the underlying and surface levels, intermediate representations are mere epiphenomena of sequential rule application. This usage may, however, be taken as supporting Goldsmith's general point that a strong emphasis on derivations at certain points in the historical development of the field may obscure the importance of levels.⁶

Let us now illustrate each of the rule interactions in (1) in turn. The data in (2) show alternations between the plural and neuter singular forms of adjectives in Norwegian (cf. Kristoffersen, 2000: 109). The plural is marked by

a suffix {-ə}, the neuter singular by {-t}. In a trilateral cluster of Son[◌]Obs[◌]Obs, the first obstruent is deleted (Cluster Simplification), as shown in the neuter singular forms for ‘dull’ and ‘glossy’. The neuter singular form of ‘clever’ shows a second alternation. A coronal consonant coalesces with a preceding /r/ to yield the corresponding retroflex. Here, the sequence /rt/ coalesces to [t]. As the neuter singular forms of ‘sharp’ and ‘dark’ show, Cluster Simplification must precede, and feed, Coalescence. UR = underlying representation; SR = surface representation.

(2) UR	PL	N.SG	
/tyn/	tynə	tynt	‘thin’
/kvalm/	kvalmə	kvalmt	‘nauseous’
/læ:r/	læ:rə	læ:t	‘clever’
/sva:r/	sva:rə	sva:t	‘huge’
/blaŋk/	blaŋkə	blaŋt	‘glossy’
/skarp/	skarpə	skat	‘sharp’
/mørk/	mørkə	mœt	‘dark’

Sample derivations for the neuter singular forms of ‘clever’, ‘glossy’, and ‘sharp’ are given in (3). Where the structural description is not met at the relevant point in the derivation, we write *n/a* (not applicable).

(3) UR	/læ:r + t/	/blaŋk + t/	/skarp + t/
Cluster Simplification	<i>n/a</i>	blaŋt	skart
Coalescence	læ:t	<i>n/a</i>	skat
SR	[læ:t]	[blaŋt]	[skat]

An example of bleeding is provided by Lithuanian (Kenstowicz and Kisseberth, 1973: 2; see also Ambrazas, 1997: 75).

(4) UR	INF	1SG.FUT	IMP.PL	
/ger-/	gerti	gersiu	gerkite	‘drink’
/kas-/	kasti	kasiu	kaskite	‘dig’
/tek-/	tekeeti	teksiu	tekite	‘flow’
/dresk-/	dreksti	dreksiu	drekskite	‘bind’

The alternations in (4) reflect two operations: Metathesis, which exchanges a fricative and a velar stop before another consonant, and Degemination. Metathesis bleeds the application of Degemination in the imperative plural of

'bind', but feeds it in the first person singular future form. Derivations for 'to bind', 'flow (IMP.PL)' and 'bind (IMP.PL)' are shown in (5).

(5) UR	/dresk + ti/	/tek + kite/	/dresk + kite/
Metathesis	dreksti	n/a	drekskite
Degemination	n/a	tekite	n/a (<i>bled</i>)
SR	[dreksti]	[tekite]	[drekskite]

Turning to counterfeeding, it is useful following McCarthy (1999) to distinguish **counterfeeding on the environment** from **counterfeeding on the focus**. Counterfeeding on the environment may be illustrated by the interaction of Post-sonorant Voicing and Syncope in Tangale, a Chadic language spoken in Nigeria described by Kidda (1985, 1993), and Kenstowicz and Kidda (1987). Examples are given in (6). Vowels with an underdot are [RTR]. Suffixal vowels harmonize with root vowels in this feature.

(6) UR	SG	DEF.SG	1SG.POSS	
/lɔɔ/	lɔɔ	lɔɔ-í	lɔɔ-nó	'meat'
/bɔgat/	bɔgat	bɔgat-í	bɔgad-nó	'window'
/adɔk/	adɔk	adɔk-í	adɔg-nó	'load'
/wudó/	wudó	wud-í	wud-nó	'tooth'
/tɔʊʒɛ/	tɔʊʒɛ	tɔʊʒ-í	tɔʊʒ-nó	'horse'
/lútu/	lútu	lút-í	lút-nó, *lúd-nó	'bag'
/dɔka/	dɔka	dɔk-í	dɔk-nó, *dɔg-nó	'salt'

Preceding a sonorant consonant, stops surface as voiced. This is seen in the first person singular possessive forms for 'window' and 'load'. A stem-final short vowel is elided when any phonological material follows (Kenstowicz, 1994: 96), as shown in all suffixed forms of 'tooth' and 'horse'. Elision fails to feed Pre-sonorant Voicing. Derivations of the first person possessive forms of 'window', 'horse' and 'bag' are given in (7).

(7) UR	/bɔgat+nó/	/tɔʊʒɛ+nó/	/lútu+nó/
Pre-sonorant Voicing	bɔgadnó	n/a	n/a (<i>counterfed</i>)
Syncope	n/a	tɔʊʒnó	lútnó
SR	[bɔgadnó]	[tɔʊʒnó]	[lútnó]

Counterfeeding on the focus refers to **chain shift** phenomena. An example is furnished by raising in Nzebi, a Bantu language of Gabon (Guthrie, 1968; Clements, 1991; Kirchner, 1996). Before certain suffixes containing /i/, a vowel is raised, but only to the next level up, as illustrated in (8). In other words, raising

does not feed further raising. In rule-based terms, raising from mid to high must precede raising from low to mid.

(8) UR

/bis/	bis	bisi	'to refuse'
/suem/	suem	suemi	'to hide oneself'
/kolən/	kolən	kulini	'to go down'
/bet/	bet	biti	'to carry'
/βeəd/	βeəd	βeedi	'to give'
/tɔəd/	tɔəd	toodi	'to arrive'
/sal/	sal	seli	'to work'

The classic example of counterbleeding comes from Canadian English, as given in (9) (Joos, 1942; Chambers, 1973). According to the standard account, the diphthongs /āy/ and /āw/ are raised to [ʌy] and [ʌw] before voiceless obstruents. Between a vowel and an unstressed vowel, the coronal stops /t d/ are both realized as a flap [ɾ]. To derive forms such as *writing*, however, Raising must apply before Flapping.

(9) UR

/wāyp/	wʌyp	wʌypɪŋ	<i>wipe</i>
/hāyk/	hʌyk	hʌykɪŋ	<i>hike</i>
/wed/	wed	werɪŋ	<i>wade</i>
/ʃut/	ʃʊt	ʃʊrɪŋ	<i>shoot</i>
/rāyt/	rʌyk	rʌyɪŋ	<i>write</i>
/rāyd/	rāyd	rāyɪŋ	<i>ride</i>

Example (10) gives derivations for *write*, *wading* and *writing*.

(10) UR	/rāyt/	/wed+ɪŋ/	/rāyt+ɪŋ/
Raising	rʌyt	<i>n/a</i>	rʌytɪŋ
Flapping	<i>n/a</i>	werɪŋ	rʌyɪŋ
SR	[rʌyt]	[werɪŋ]	[rʌyɪŋ]

Up to this point we have considered the four possible two-rule interactions. There are also crucial three-rule interactions described in the literature. One of these is **double counterbleeding**. The best known example is Yawelmani Yokuts (Newman, 1944; McCarthy, 1999). This language has a process of Vowel Harmony

that spreads lip rounding from the stem onto a suffixal vowel of the same vowel height, for example, *dub-hun* 'lead by the hand -NONFUT' (cf. *xil-hin* 'tangle-NONFUT'; *bok'-hin* 'find-NONFUT'). This is counterbled by a process that lowers all underlying long high vowels to their mid counterparts. Lowering is then itself counterbled by a rule that shortens vowels in a closed syllable, for example, *ʃap-al* 'burn-DUBITATIVE', but *ʃap-hin* 'burn-NONFUT'. The effect of these rules on an underlying form like /ʔu:ʔ-hin/gives [ʔoʔhun].

Other kinds of three-rule interactions combine feeding/bleeding with counterfeeding/counterbleeding. Two kinds of interaction that have played an important part in recent debate are known by special names: **rule sandwiching** and **Duke-of-York derivations**. The reason for singling out these two will become clear when we look in more detail at OT.

Rule sandwiching (Levi, 2000; Bye, 2001) refers to three-rule interactions of the form $\mathcal{P} > \mathcal{Q} > \mathcal{R}$, in which (i) rules \mathcal{P} and \mathcal{R} introduce identical structural changes, and (ii) \mathcal{P} feeds/bleeds (i.e. interacts transparently with) \mathcal{Q} , but \mathcal{R} counterfeeds/counterbleeds \mathcal{Q} . Here we will illustrate the phenomenon with data from Mohawk, a Northern Iroquoian language spoken in New York State, Ontario and Quebec. The examples are in (11) taken from Hagstrom (1997, 114ff.). In this language, epenthetic vowels (shown underlined) may or may not be visible to stress assignment. The metrical visibility of epenthetic vowels is a function of the phonotactic conditions which trigger the epenthesis.

- | | | |
|----------------------------|------------------|----------------------------------|
| (11) /wak-haratat-u-hatye/ | wakharatatuhátye | 'I go along lifting up' |
| /k-ataʔkerahkw-haʔ/ | kataʔkerákwaʔ | 'I float' |
| /ʌ-k-r-ʌ-ʔ/ | ákerʌʔ | 'I will put it into a container' |
| /waʔ-t-k-atat-nak-ʔ/ | waʔkatátɛnakeʔ | 'I scratched myself' |
| /wak-nyak-s/ | wakényaks | 'I get married' |
| /te-k-ahsutr-haʔ/ | tekahsutérhaʔ | 'I splice it' |

Canonical stress falls on the penultimate syllable, as shown in the forms meaning 'I go along lifting it up', and 'I float'. Mohawk has two rules that epenthesize *e*, one that operates in trilateral consonant clusters, and another that breaks up biliteral consonant clusters. Epenthesis into biliteral clusters is invisible to stress. Epenthesizing once thus shifts the stress one position to the left of its canonical position. Thus, in the form meaning 'I will put it into a container', the stress is displaced to the antepenultimate syllable since the epenthetic vowel surfaces in the penultimate position. In 'I scratched myself', biliteral epenthesis has occurred twice, resulting in a second stress shift, with stress surfacing in the pre-antepenultimate syllable. Epenthesis into a trilateral cluster, however, feeds *into* stress placement, and stress surfaces in its canonical position. This is shown by the last two forms, meaning 'I get married' and 'I splice it'. Stress

Assignment is thus apparently ‘sandwiched’ between two distinct epenthesis rules. (12) shows how this difference is derived in terms of an ordered sequence of rules.

(12) UR	/Λ-k-r-Λ-ʔ/	/wak-nyak-s/
Triconsonantal Epenthesis	<i>n/a</i>	wakenyaks
Stress	ákrΛʔ	wakényaks
Biconsonantal Epenthesis	ákerΛʔ	<i>n/a</i>
SR	ákerΛʔ	wakényaks

Duke-of-York derivations (Pullum, 1976) are three-rule interactions of the form $\mathcal{P} > \mathcal{Q} > \mathcal{R}$, where \mathcal{P} creates some string \mathcal{S} that feeds \mathcal{Q} , and \mathcal{R} counterbleeds \mathcal{Q} by destroying the very string \mathcal{S} that \mathcal{P} created.⁷ It can be illustrated by Canadian French (Poliquin, 2006), data from which are shown in (13).

(13) UR			
/pœti/	pœ.tʰi	<i>Petit</i>	‘small’
/rabru/	ʁa.bʁu	<i>rabroue</i>	‘scold’
/apik/	a.pik	<i>à-pic</i>	‘cliff’
/ekrul/	e.kʁul	<i>écroule</i>	‘crumble’
/filip/	fi.lip	<i>Philippe</i>	name
/tymuʃ/	tʰy.muʃ	<i>tue-mouche</i>	‘fly paper’
/elev/	e.le:v	<i>élève</i>	‘pupil’
/elɔʒ/	e.lɔ:ʒ	<i>éloge</i>	‘praise’
/egliz/	e.gli:z	<i>église</i>	‘church’
/misiv/	mi.si:v	<i>missive</i>	‘letter’
/difyz/	dʰi.fy:z	<i>diffuse</i>	‘diffuse’

In word-final open syllables, high vowels are tense, but in word-final closed syllables, high vowels are obligatorily lax, as can be seen from comparing the forms for ‘small’ and ‘scold’ with those for ‘steep’ and ‘crumble’. According to Poliquin’s interpretation, Final Closed Syllable Laxing feeds an optional rule of Vowel Harmony, which additionally introduces lax vowels into non-final open syllables. This is illustrated by the forms for ‘Philippe’ and ‘fly paper’. Vowels are lengthened preceding coda voiced fricatives (‘pupil’, ‘praise’). Now, lengthening does not generally give a tense vowel, but lengthened high vowels are always [+ATR], as in ‘church’. This is the crucial Duke-of-York U-turn, since it gives rise to surface disharmonic forms like *mi.si:v* and *dʰi.fy:z*. This, Poliquin

argues, points to an intermediate representation **mi.siv* where Final Closed Syllable Laxing has applied, followed by Harmony. Laxing is subsequently reversed by Pre-voiced fricative Tensing. Derivations are given in (14) for *à-pic*, *Philippe*, *élève*, and *missive*.

(14) UR	/apik/	/filip/	/elev/	/misiv/
Final Closed Syllable Laxing	a.pɪk	fi.lɪp	<i>n/a</i>	mi.siv
Vowel Harmony	<i>n/a</i>	fi.lɪp	<i>n/a</i>	mi.siv
Pre-voiced fricative Tensing	<i>n/a</i>	<i>n/a</i>	e.lɛ:v	mi.si:v
SR	[a.pɪk]	[fi.lɪp]	[e.lɛ:v]	[mi.si:v]

2.2 Depth of Derivation and Abstractness

Ideally every morpheme has a unique underlying representation. Rules allow us to minimize allomorphy by positing **abstract** underlying representations. This ideal breaks down in cases of **suppletion**, where two allomorphs are phonologically too distant to be credibly related by rule to the same underlying representation, as is the case with *go* and *went* in English.⁸

A classic demonstration of the need for abstract underlying representations, illustrated in (15), comes from Schane's 1974 discussion of Palauan, a Malayo-Polynesian language.⁹ Stress falls on the last vowel of the stem when there is a suffix (in both the conservative and innovative future participle forms), otherwise on the penultimate vowel (in the present middle form). Unstressed vowels are realized as schwa.

(15) PRES.MIDDLE	FUT.PART (conservative)	FUT.PART (innovative)	UR	
mədánəb	dəŋəbáll	dəŋóbəl	/dɑŋɔb/	'cover'
mətéʔəb	təʔəbáll	təʔíbəl	/teʔib/	'pull out'

The stem of the verb for 'cover' has three surface allomorphs, [dánəb], [dəŋób], and [dəŋəb]. Selecting one of these surface allomorphs as the basic variant and attempting to derive the others from it will run in to problems for reasons that become clear when we consider the paradigm for 'pull out', which has the surface allomorphs [téʔəb], [təʔíb], and [təʔəb]. Vowel quality in unstressed syllables is predictable: it is always schwa. There is no similar generalization we can make with respect to vowel quality in the stressed syllable, however. The second vowels in the stems for 'cover' and 'pull out' must be specified, respectively, in the lexicon as /o/ and /i/. Since each occurrence of schwa can be

understood as the result of applying a rule of vowel reduction, we can conclude that there is an abstract underlying form, /daŋɔb/, which can be found by undoing its effects. Such examples are far from uncommon.

The same logic of reversing processes to arrive at the underlying form invites the possibility of positing segments that are never realized on the surface. Kiparsky (1968 [1973]: 14) dubs this **absolute neutralization**, to distinguish it from the less controversial **contextual neutralization** where an underlying contrast is neutralized in some environments and preserved elsewhere. Absolute neutralization is a problem for acquisition because the learner is unable to appeal to surface morphophonological alternations to reconstruct the underlying representation of the segment; their only recourse is to the rules. On the assumption that rules of absolute neutralization were impossible or difficult to learn, Kiparsky proposed a stipulation, the Alternation Condition, which specified that 'Neutralization processes cannot apply to all occurrences of a morpheme'. This was later revised to read 'Neutralization processes apply only to derived forms' (Kiparsky, 1976). Nevertheless, several researchers proposed analyses in which they argued that absolute neutralization could allow considerable simplification of the grammar. A well-supported example is Hyman (1970), who proposed that Nupe, a Kwa language of Central Nigeria, has a rule lowering underlying mid vowels /ɛ/ and /ɔ/ to /a/ in all environments. The evidence for this abstract solution comes in part from the distribution of palatalized and labialized consonants. With one exception, palatalized consonants are found exclusively preceding front unrounded vowels, and labialized consonants before back rounded vowels. This distribution is thus consistent with being determined by rule. Preceding a low vowel /a/, however, a consonant may be plain, palatalized or labialized. Examples are shown in (16).

(16) UR

/ɛgũ̃/	ɛg ^w ũ̃	'mud'
/ɛgó/	ɛg ^w ó	'grass'
/ɛgĩ/	ɛg ^y ĩ	'child'
/ɛgē/	ɛg ^y ē	'beer'
/ɛgā/	ɛgā	'stranger'
/ɛgɔ̃/	ɛg ^w ā	'hand'
/ɛgè/	ɛg ^y ā	'blood'

Hyman accounts for the distribution of secondary articulation by allowing lowering of the low mid vowels /ɛ ɔ/ to /a/ to counterbleed allophonic palatalization/labialization preceding a non-low vowel. He also cites evidence in favour of this analysis from the adaptation of Yoruba loanwords in Nupe, for example, Yoruba *t̄r̄ɛ* > Nupe *t̄^wār^yā* 'to give a gift'. See Kiparsky (1976) for an overview of other similar proposals.

2.3 Natural and Unnatural Rule Orders

Early generative phonology had a preoccupation with naturalness. This was obviously true in the development of distinctive feature theory and, to varying degrees, the search for formal mechanisms, such as non-linear representational frameworks (such as Autosegmental Phonology, Feature Geometry and Metrical Phonology), that would bring greater elegance and simplicity to the statement of natural rules. However, this same interest also disclosed itself in attempts to discern whether certain ordering relationships between rules could be said to be more natural or ‘unmarked’ than others. Related to the problem of absolute neutralization, is an argument by Schane (1968), who considers the interaction of Vowel Nasalization and Nasal Deletion in French. Nasal Deletion counterbleeds Vowel Nasalization, as in (17), which shows the derivations for *bon* ‘good’ and *bonté* ‘goodness’.

(17) UR	/bɔn/	/bɔnte/
Vowel Nasalization	bɔ̃n	bɔ̃nte
Nasal Deletion	bɔ̃	bɔ̃te
SR	[bɔ̃]	[bɔ̃te]

Applying the same rules in the reverse order would, in addition to giving us the wrong result, destroy all the inputs to Vowel Nasalization, something that Schane calls **absolute bleeding**. Schane argues that orderings that would give this result are universally impossible.

Kiparsky (1968: 200) proposes what we might call the **maximal application hypothesis**: ‘Rules tend to shift into the order which allows their fullest utilization in the grammar’. Recalling the definitions in (1), feeding and counterbleeding both maximize rule application since, in each case, both \mathcal{P} and \mathcal{Q} apply. See also Kenstowicz and Kisseberth (1977: 159–168) for extensive discussion of Kiparsky’s proposal.

Arguing against his earlier proposal, Kiparsky (1971) offers an alternative naturalness metric that has surface distributions as its point of departure rather than the rule interactions themselves. He proposes (p. 623) that ‘rules tend to be ordered so as to become maximally transparent’. The intuition here is that, since the condition for **transparent** rules is present in the surface representation, this makes them less complex for learners to acquire. When the conditions are absent, the rule is said to be **opaque**. The following definition in (18) is adapted from Kiparsky (1973b: 79). (For comparison, see also Kiparsky (1976: 179) and McCarthy (2007a: 108).)

- (18) A process \mathcal{P} of the form $A \rightarrow B / C_D$ is opaque if there are surface forms with any of the following characteristics:
- instances of A in the environment C_D
 - instances of B derived by the process \mathcal{P} that occur in environments other than C_D

Since McCarthy (1999) it has also been customary to talk about ‘transparent’ and ‘opaque’ interactions. Feeding and bleeding are commonly grouped together as transparent interactions, counterfeeding and counterbleeding as opaque interactions.¹⁰ OT has brought with it several additional terms for talking about opacity. McCarthy and Prince (1995, 1999) adapts the terms **underapplication** (Type 1 opacity; (18-a)) and **overapplication** (Type 2 opacity; (18-b)) to refer, respectively, to the patterns derived in rule-based approaches by counterfeeding and counterbleeding (generically, **misapplication**).¹¹ As we shall see in the next section, however, there may be reasons for severing the connection between rule interaction type and application type.

The idea that opaque rules were dispreferred was subsequently developed by Kisseberth (1973a) and Kenstowicz and Kisseberth (1973). Kenstowicz and Kisseberth (1973, 1977) argue specifically that the bleeding order should be considered unmarked in interactions between certain kinds of rules. For example, cross-linguistically epenthesis overwhelmingly bleeds voicing assimilation. The reverse ordering is to date only known from one language, New Julfa Armenian (Vaux, 1998).

Recently, there has been a renewal of interest in the question of natural rule ordering. A survey by Parker (2009) of rule interactions in seven languages finds clear preferences for transparent interactions, with feeding and bleeding, respectively, making up 54.8 per cent and 24.4 per cent of observed tokens. The incidence of counterbleeding and counterfeeding, at 11.8 per cent and 9.0 per cent is considerably lower. Ettliger (2009), however, who specifically looks at the interaction of vowel harmony with other rules, finds that counterfeeding is very common, while bleeding is very rare.

A possible alternative to extrinsic ordering debated in the 1970s was the **global rule** (Kenstowicz and Kisseberth, 1970; Kisseberth, 1973a and b; Dinnsen, 1974; Kenstowicz and Kisseberth, 1977), a.k.a. **derivational constraint**. In standard derivational phonology, rules apply in strictly Markovian fashion to the current representation with no possibility of looking back to previous states or forward to optimal future states. A global reinterpretation of the *writer* versus *rider* contrast, for example, might reformulate the Raising rule so that it only applies before underlyingly voiceless consonants. The forward-looking type of global rule was motivated by the idea that rules sharing a functional teleology could enter into **conspiracies** (Kisseberth, 1970a and b). Ultimately, global rules

were rejected as having unwarranted generative power (see, for example, Kiparsky, 1976), although many of the issues would return in the 1990s with the advent of constraint-based approaches to phonology. The two closely related approaches described by Goldsmith (1993) and Lakoff (1993) propose rules that are sensitive to underlying level of representation. This conception also made it into the earliest attempt to deal with the problem of opacity in OT (McCarthy, 1996), which also forms part of the basis of Diagonal Correspondence Theory (Ettlinger, 2009).

2.4 Relations between Rule Interaction Type and Application Type

In the previous section we suggested that there was a straightforward relation between rule interaction type and application type: feeding and bleeding translate into normal application, counterfeeding into underapplication, and counterbleeding into overapplication. Recent papers by Baković (2007, to appear), Anttila et al. (2008) and Kavitskaya and Staroverov (2009), however, argue in one way or another for breaking this connection. A typology of the attested combinations of interaction and application type is provided by Baković (to appear).

Kavitskaya and Staroverov (2009) illustrate a type of rule interaction they dub **fed counterfeeding**, where \mathcal{P} feeds \mathcal{Q} and \mathcal{Q} counterfeeds \mathcal{P} . In Lardil (Pama-Nyungan, Queensland; for example, Hale, 1973), the absolutive forms of nouns whose stem is longer than two moras, a final vowel is deleted. Apocope feeds a rule that deletes final non-coronal consonants, giving derivations such as the following /dibirdibi/→[dibirdib]→[dibirdi] ‘rock cod’. In derivations like this one, Consonant deletion counterfeeds Apocope. The fact that Apocope only applies in the absolutive form, however, raises the question whether this interpretation of the facts is the correct one. An alternative analysis in which the deletion of the final vowel was due to a morphological subtraction process, as assumed, for example, by Horwood (2001) and Kurisu (2001), would have to be carefully ruled out.

Anttila et al. (2008) discuss several processes affecting consonant clusters in Singapore English and an interaction Baković terms **surface-true counterfeeding**. Sibilant clusters are broken up by epenthesis, for example, /reiz+z/→[reizəz] *raises* (cf. [bægz] *bags*), and the second of two obstruents in the coda is deleted, for example, /test/→[tes] *test* (cf. [testɪŋ] *testing*). Deletion thus creates potential inputs to Epenthesis, for example, /list+z/→[lisz], but instead of feeding Epenthesis, giving counterfactual *[lisəz], we find that Deletion feeds a rule of Degemination: /list+z/→[lisz]→[lis]. As Baković (to appear) points out, Epenthesis is counterfed by Deletion, but it does not underapply since

Degemination eliminates the input for Epenthesis, with the result that Epenthesis remains surface-true.

Baković further argues that certain types of feeding may result in opacity. Some alternations suggest **self-destructive feeding**, where \mathcal{P} feeds \mathcal{Q} and \mathcal{Q} counterbleeds \mathcal{P} . The example Baković provides is from Turkish, where a putative rule that deletes [s] following a consonant feeds (and is counterbled by) a second rule eliding a [k] between vowels, for example, /ajak+su/ → [ajaku] → [aja.u] ‘his foot’ (cf. [oda-su] ‘his room’, /tʃan+su/ → [tʃanu] ‘his bell’, /bebek+i/ → [bebe.i] ‘baby-ACC’). The alternation between /s/ and \emptyset is limited to the genitive and third person possessive suffixes, so it appears we may be dealing with two allomorphs {-sI} (after vowels) and {-I} (after consonants) rather than a phonological rule. What is nevertheless interesting is that {-I} is still selected by consonant-final stems that lose their final consonant (opaque allomorph selection). Other s-initial suffixes do not evince the alternation, for example, [fajda-suz] ‘without use’, [akul-suz] ‘without intelligence, stupid’.

Another type of opaque feeding comes from Classical Arabic, where words beginning with complex onsets prothesize a vowel. Since vowel-initial words are disallowed in Arabic, this further feeds the insertion of a glottal stop, for example, /ktub/ → [uktub] → [ʔuktub] ‘write! (M.SG)’. This type of interaction, which Baković calls **non-gratuitous feeding**, entails a kind of overapplication: glottal epenthesis separates the epenthetic vowel from the word boundary that was part of the context of the rule, hence prothesis overapplies.

A case of **mutual bleeding**, or **non-opaque counterbleeding**, is the interaction of Epenthesis and Elision in Lardil. Clusters of $i\bar{u}$ are broken up by insertion of a glide [w], for example, /papi+u/ → [papiwu] ‘father’s mother -ACC.FUT’. In general, however, vowel clusters undergo simplification through deletion of the second vowel, for example, /tʰæmpæ+u/ → [tʰæmpæ] ‘mother’s father -ACC.FUT’. Epenthesis bleeds Elision because it eliminates the vowel hiatus. It is also the case that Elision counterbleeds Epenthesis, but Epenthesis nevertheless cannot be said to overapply since Elision does not actually destroy the context for Epenthesis.

Baković’s most controversial proposal is **cross-derivational feeding**. Considerations of space preclude going into detail here, although see endnote 13.

2.5 No Rule Ordering and Persistent Rules

The preceding sections have looked at determinate sequential ordering, where rules apply in a strict sequence. In this section, we look at theories where the grammar does not specify an ordering. Koutsoudas (1976) distinguishes several possible conventions for rule application where no ordering is specified. Here we will concentrate on two, **random sequential ordering** and **simultaneous**

application. These are still of relevance today also because of their similarities to serial and parallel implementations of OT.

In random sequential ordering, a rule or process applies whenever its structural description arises in the course of a derivation. The same structural description may arise more than once, and in this case the rule applies each time its input occurs. A rule with this property is known as **persistent** (Chafe, 1968).¹² As already pointed out, rule persistence was a feature of the theory of rule interaction advanced by Koutsoudas et al. (1974 [1971]). It was also incorporated into Natural Generative Phonology (Vennemann, 1972b; Hooper, 1976). In a grammar allowing only random sequential ordering, rules may only interact by feeding since, if \mathcal{P} creates an input to \mathcal{Q} , then \mathcal{Q} must apply – the grammar does not have to specify the order $\mathcal{P} > \mathcal{Q}$. In cases of putative bleeding, however, the structural descriptions of both \mathcal{P} and \mathcal{Q} are simultaneously present in the input representation, requiring an arbitrary stipulation that \mathcal{P} applies first. In this connection Hooper (1976) discusses the interaction of epenthesis and voicing assimilation in English. In brief, the English plural suffix {-z} assimilates in voicing to a preceding voiceless obstruent ($/kæt+z/ \rightarrow [kæts]$). When the stem ends in a sibilant, a vowel is inserted ($/meiz+z/ \rightarrow [meizɪz]$). But when the stem-final sibilant is voiceless, it is the epenthesis rule that has priority ($/meis+z/ \rightarrow [meisɪz], *[meisis]$). At best, random sequential application predicts that we should find variation between forms that are the outcome of bleeding interaction ($[meisɪz]$) and those that result from counterbleeding ($*[meisis]$). Hooper (p. 73f.) is therefore forced to adopt a solution in which the plural suffix is lexically specified as a disjunction of two allomorphs, {-ɪz} after sibilants, and {-z} elsewhere. If it was merely a matter of describing the variation in shape of a particular morpheme, this allomorphic solution would not be particularly costly. Unfortunately, it ends up missing several generalizations. For one thing, epenthesis in English seems to be a general process that applies whenever the suffix is sufficiently similar to the stem-final consonant. The third person singular suffix thus behaves in exactly the same way as the plural, and the past tense suffix {-d} behaves analogously following a coronal stop ($/pæt+d/ \rightarrow [pætɪd]; /pæd+d/ \rightarrow [pædɪd], *[pættɪ]$). Without recourse to bleeding we would be forced to duplicate the allomorphic solution for each suffix. Second, as observed by Kenstowicz and Kisseberth (1977), epenthesis overwhelmingly bleeds voicing assimilation cross-linguistically, which would force us to duplicate the allomorphic solution across languages. Finally, the allomorphic solution could not deal with alternations in stems, since these pertain to open lexical classes, such as noun, adjective and verb.¹³

For example, we observed above that Metathesis may bleed Degemination in Lithuanian. Positing listed allomorphs for each stem that participates in the alternation (e.g. $\{dresk-\} \sim \{dreks-\}$ 'bind') would not help deal with previously unencountered stems ending in $s\check{c}k$. Government Phonology (Kaye, 1992, 1995) also

takes the line that all rules are persistent (Kaye, 1992: 141), and so similarly rules out the possibility of bleeding. However, this approach places great emphasis on representations in explaining phonological phenomena, which takes much of the descriptive brunt off rule interaction. Epenthesis, for example, is not dealt with in terms of processes, but by conditions on the phonetic interpretation of phonological structures, in which case the issue of ordering with respect to epenthesis at least does not arise (see Section 3 for an example).

Counterfeeding and counterbleeding are also incompatible with the logic of random sequential application that rules apply whenever their structural description is met. In counterfeeding, a rule \mathcal{P} fails to feed another rule \mathcal{Q} whose structural description it creates. Counterbleeding is problematic for the same reason as bleeding. As we shall see in Section 3, there has been much research in OT and Government Phonology to try and reinterpret opaque effects representationally or as non-phonological.¹⁴

The basic premise of simultaneous rule application is that the underlying representation is the sole determinant of rule function. This entails that the application of a rule is not affected by the application of any other rule – there is no rule interaction. As Kenstowicz and Kisseberth (1979: 291–307) show, this entails that simultaneous rule application can only deal with counterbleeding and counterfeeding interactions, but not feeding or bleeding interactions.

Many of the issues regarding simultaneous and random sequential application we have been discussing return in a different guise in Optimality Theory (OT) (Prince and Smolensky, 2004 [1993]; McCarthy and Prince, 1993b). We will not go into details about how OT works here. For an introduction, see Uffmann (this volume). Here we will simply note some core properties of Classical parallelist OT that pose special problems for dealing with certain types of interaction.

First, in Classical OT, all processes apply simultaneously (in parallel) to the input form. However, this does not imply that OT and simultaneous rule application fare similarly with respect to capturing the two-rule interactions listed in (1) since simultaneous application is **input-oriented**, while OT is **output-oriented**. For this reason, simultaneous rule application can capture opacity effects, but OT cannot because the Faithfulness violation that results from the opacified process is not motivated by the surface structure. As a consequence of gratuitous violation of some Faithfulness constraint, a candidate representation can_{op} corresponding to the output of an opaque interaction will always be **harmonically bounded** by a more optimal transparent candidate can_{trans} with a proper subset of can_{op} 's violation marks. In simultaneous rule application, on the other hand, the condition for the application of the process of course lies in the input to the rule. Solutions to the opacity problem in OT generally frame the issue as a **ranking paradox**. The desired winner can_{op} can

be selected if there is some highly ranked constraint that *cand_{op}* satisfies and its transparent competitor violates.

More recently, there has been increasing interest in a version of OT now known as Harmonic Serialism (HS-OT). HS-OT was first aired in Prince and Smolensky (2004 [1993]: 94–95), but serious interest in exploring the possibilities of this implementation is very recent (McCarthy, 2008a and b). In HS-OT, processes apply in a random sequential order.¹⁵ GEN is limited to making a single **harmonic improvement** at a time (**gradualness**). Leaving aside for now the somewhat difficult problem of what should count as a single change, EVAL then determines the locally optimal output from the small set of candidates on the current pass, and the output is fed back to GEN, which then produces another set of candidates for evaluation. This process continues to loop until the latest output of EVAL is identical to the latest input to GEN – if no further harmonically improving single edits can be made, the derivation is said to ‘converge’. HS-OT’s serialism is of little help from the point of view of solving the opacity problem, however, since processes apply in a random order. At the very least what is required is some formal mechanism which would allow the OT grammar to mimic the extrinsic ordering of processes. Regardless whether the implementation is parallel or serial, developing the necessary mechanism is something of a technical challenge, although it has been attempted with some success for both parallel OT (Sympathy Theory; McCarthy, 1999) and HS-OT (Candidate Chains Theory, or OT-CC; McCarthy, 2006). Sympathy and OT-CC have been shown to be able to describe two-rule interactions, as well as multiple opacity as in Yawelmani Yokuts. But for quite fundamental reasons, these attempts may represent the horizon of what opaque interactions it is possible to describe in OT. Specifically, even if we allow the linear ordering of *processes*, there would seem to be no way to allow the OT grammar to mimic linear ordering of *rules*, which would require a pairing of process and context. Recall that processes in GEN are precisely context-free. Reintegrating processes with the contexts in which they apply would amount to abandoning OT.¹⁶ The implications of this are particularly apparent in the case of rule sandwiching, since what is at issue here is precisely getting the same process to apply at different points depending on context. Recall that in Mohawk, epenthesis of *e* must apply crucially earlier in trilateral clusters than in biliteral clusters.

Another strategy for OT in dealing with opacity is to identify some function that opacity serves, allowing a solution to the constraint ranking paradox. Kaye (1974) argued that opacity should be considered natural, noting that it permits recovery of underlying contrasts. For example, the presence of the raised allophone in Canadian English [rʌyɾə], signals that the flap is ‘really’ a voiceless obstruent, thereby providing an important clue for identification of the word. The idea that opacity may be functional in this way, as opposed to a mere epiphenomenon of sequential rule application, inspires pretty much every

attempt to deal with opacity from the perspective of parallel OT (McCarthy, 1999, 2006; Goldrick, 2000; Łubowicz, 2003; Ettliger, 2009). This also inherently limits the kinds of opaque interactions that recoverability-based approaches to opacity can deal with. Specifically, opaque contexts that are not present in underlying representation but arise through the application of processes, such as occurs in Duke-of-York derivations, cannot be dealt with (Idsardi, 1997). This is because the recovery of non-underlying opaque contexts would by definition not contribute to recovering the underlying form. In sum, abstracting away from the details of any one implementation of OT, rule sandwiching and Duke-of-York derivations pose problems to both parallel and harmonic serial OT for reasons which have to do with the fundamental character of the theory. For this reason, the reality of both is a hotly contested area of contemporary debate over the relative merits of OT and rule-based phonology (e.g. Vaux, 2008). McCarthy (1999, 2006) argues that genuine Duke-of-York effects do not exist, while other researchers, including Idsardi (2000), Poliquin (2006) and Odden (2008), have argued that they represent real phenomena.

Let us see how each of the two-rule interactions described in Section 2.1 fare in HS-OT.¹⁷ A corresponding OT analysis of the Norwegian feeding case presented in (2) might be constructed using the constraints in (19).

- (19) a. *CCC
A cluster of three consonants is disallowed.
b. MAX-C
A consonant in the input must have a correspondent in the output.
c. * r^{\wedge} Cor
A sequence of /r/ followed by a coronal consonant is disallowed.
d. UNIFORMITY
A segment in the output must not have multiple correspondents in the input.

*CCC must outrank MAX-C and * r^{\wedge} Cor outrank UNIFORMITY. Even though the Markedness constraints *CCC and * r^{\wedge} Cor do not conflict, the logic of Harmonic Serialism additionally requires that the former outrank the latter. If they were unranked, the mapping from /skarpt/ to intermediate [skart] would not constitute a harmonic improvement: simplifying the cluster would merely trade one violation mark for another that the grammar deemed equally bad.

A serial evaluation is shown in Tableau 6.1. On the first pass, the locally optimal candidate satisfies *CCC at the expense of incurring a lower-ranked violation of MAX-C, but it still violates lower ranked * r^{\wedge} Cor. On the second pass, the locally optimal candidate from the first pass is bested by candidate (b), which additionally satisfies * r^{\wedge} Cor, gaining a mark on lower-ranked UNIFORMITY.

Tableau 6.1

a. *First Pass*

/skar ₁ p ₂ + t ₃ /	*CCC	*r̄Cor	MAX-C	UNIFORMITY
a. skar ₁ p ₂ t ₃	*!			
b. \Rightarrow skar ₁ t ₃		*	*	

b. *Second Pass*

/skar ₁ p ₂ + t ₃ /	*CCC	*r̄Cor	MAX-C	UNIFORMITY
a. skar ₁ t ₃		*!	*	
b. \Rightarrow skat _{1,3}			*	*

Turning to the bleeding interaction of Metathesis and Degemination in Lithuanian (4), at least the following additional constraints in (20) would be needed.

(20) a. *s̄k̄C

A cluster of s̄k̄ followed by a consonant is disallowed.

b. *GEMINATE

Geminates are disallowed.

c. LINEARITY

The output is consistent with the precedence structure of the input.

Tableau 6.2 illustrates the serial evaluation. The Markedness constraints *s̄k̄C and *GEMINATE both receive violation marks in the fully faithful candidate (a). Since the application of either Metathesis or Degemination eliminates both violation marks, the choice as to which process applies must be arbitrated entirely by the Faithfulness constraints LINEARITY and MAX-C. Ranking LINEARITY below MAX-C ensures that the candidate implementing Metathesis (c) wins over the candidate where Degemination has applied (b). On the second pass, no further harmonic improvements can be made, so the derivation converges.

Now let us address counterfeeding in Tangale (6). An OT analysis might use the constraints in (21). The first constraint ELISION is a shameless imitation of the corresponding rule; hardly a serious candidate for inclusion in the universal constraint set CON. It is nevertheless adopted here as an expository convenience.

Tableau 6.2

a. *First Pass*

/dresk+kite/	*s [~] k [~] C	*GEMINATE	MAX-C	LINEARITY
a. dreskkite	*!	*		
b. dreskite			*!	
c. dre skite				*

b. *Second Pass*

/dresk+kite/	*s [~] k [~] C	*GEMINATE	MAX-C	LINEARITY
a. dre skite				*

(21) a. ELISION

A stem-final short vowel is disallowed preceding phonological material.

b. MAX-V

A vowel in the input must have a correspondent in the output.

c. PRE-SONORANT VOICING

A sequence of voiceless obstruent followed by a sonorant is disallowed.

d. IDENT[voice]

Corresponding segments must have identical specifications for [voice].

In order to derive the effects of Elision and Pre-sonorant voicing in isolation, ELISION must outrank MAX-V and PRE-SONORANT VOICING outrank IDENT[voice]. Next we have to get the two processes to interact. Simply combining these rankings, however, incorrectly predicts that Elision should feed Pre-sonorant Voicing. In Harmonic Serialism, processes apply persistently, until no further harmonically improving changes can be made. Assuming ELISION dominates POST-SONVOI, the mapping from underlying /lútu+nó/ to the desired output form [lútnó] is harmonically improving. Unfortunately, the derivation does not converge at this point, because further harmonic improvements are possible. A second pass through GEN and EVAL converges on *[lúdnó], which does even better than the previous output and desired winner by transparently applying Pre-sonorant Voicing. In the serial evaluation shown in Tableau 6.3, the desired winner is indicated with '☞', the undesired but actual winner with '☞'. The crucial fatal mark is indicated with an inverted exclamation point '¡'.

Tableau 6.3

a. *First Pass*

/lútu+nó/	ELISION	POST-SONVOI	MAX-V	IDENT[voice]
a. lútunó	*!			
b. lútunó lútnó		*	*	

b. *Second Pass*

/lútu+nó/	ELISION	POST-SONVOI	MAX-V	IDENT[voice]
a. lútunó lútnó		i*!	*	
b. lútnó lúdnó			*	*

McCarthy (2006) proposes an implementation of OT that essentially reintroduces extrinsic ordering of processes. The formal implementation of this idea is somewhat complex, so I shall limit myself to explaining the basic idea. McCarthy reifies the derivation as a kind of representation he dubs the **candidate chain**, which is subject to harmonic evaluation by EVAL. The candidate chain is a linear ordering of forms, beginning with the fully faithful form and ending with the intended output. The chain may also contain intermediate representations. As in Harmonic Serialism, each form in the chain must represent a single harmonic improvement over the previous one. Markedness and Faithfulness constraints only evaluate the last form in the candidate chain. But by reifying the derivation in this way, it is possible to formulate constraints on relations between forms in the chain. In the Tangale case, the comparison of two chains provides the focus of interest, the undesired $\langle \text{lútunó, lútnó, lúdnó} \rangle$, and the desired $\langle \text{lútunó, lútnó} \rangle$. The chains differ in the way that violations of Faithfulness constraints are introduced, that is, in the way they mimic the extrinsic ordering of processes. In both chains there is a violation of MAX-V (the result of a process of vowel deletion) in the second form of the chain, [lútnó]. In the first chain, however, there is also a violation of IDENT[voice] (the result of a process of voicing). McCarthy proposes that the first candidate chain can be ruled out by a constraint (one of a family called PRECEDENCE) that says violations of MAX-V cannot be followed in the chain by violations of IDENT[voice]. Crucially, this constraint is violated by the transparent chain, but not by the opaque chain. When the constraint is included, the desired opaque winner is no longer harmonically bounded by the transparent candidate. Ranked at the appropriate position in the

hierarchy, the newly identified constraint permits selection of the opaque candidate as the winner.

Finally, let us see how HS-OT fares on counterbleeding opacity. An OT analysis of the interaction of Canadian Raising and Flapping from (9) may be constructed using the constraints in (22).

(22) a. RAISING

The diphthongs /āy/ and /āw/ are disallowed preceding a voiceless obstruent.

b. IDENT[low]

Corresponding segments must have identical specifications for [low].

c. *VTV

A coronal stop is disallowed between vowels.

d. IDENT[sonorant]

Corresponding segments must have identical specifications for [sonorant].

To derive the effects of Raising and Flapping in isolation, RAISING must outrank IDENT[low] and *VTV outrank IDENT[sonorant]. Again, simply combining the rankings predicts that the transparent candidate (c) should win. The flapping of /t/ satisfies both RAISING and *VTV in one fell swoop. Since the desired winner *rāyriŋ* entails two changes, it cannot be generated on the first pass. On the second pass, though, the optimal candidate is identical to the input from the previous pass, so the derivation converges incorrectly on the transparent form **rāyriŋ* as the Tableau in 6.4 show.

Tableau 6.4

a. First Pass

/rāyt+iŋ/	RAISING	*VTV	IDENT[low]	IDENT[son]
a. rāytŋ	*!	*		
b. r rāytŋ		*!	*	
c. r rāyriŋ				*

b. Second Pass

/rāyt+iŋ/	RAISING	*VTV	IDENT[low]	IDENT[son]
a. r rāyriŋ				*

The next section returns to transient rules and considers intrinsic ordering principles.

2.6 Intrinsic Rule Ordering

The 1970s saw a number of other significant attempts to restrict the power of extrinsic ordering. Intrinsic rule ordering refers to any universal principle that determines the sequential order in which rules apply, usually based on the formulation of the rules or the domain of their application (which is the topic of the next section). One of the earliest attempts at dispensing with extrinsic rule ordering was Koutsoudas, Sanders and Noll (1974 [1971]). Their Proper Inclusion Precedence was eventually rechristened as the Elsewhere Condition (Kiparsky, 1973b, 1982b, 1984), given in (23) according to Kiparsky's formulation. The Elsewhere Condition assigns priority to **disjunctively ordered** pairs of rules in the same cycle (see below).¹⁸

(23) *The Elsewhere Condition* (Kiparsky, 1984: 3)

Rules A, B in the same component apply disjunctively if and only if

- a. The input of A is a proper subset of the input of B.
- b. The outputs of A and B are distinct. In that case, A (the particular rule) is applied first, and if it takes effect, then B (the general rule) is not applied.

The Elsewhere Condition can be illustrated with data from Standard Finnish given in (24) (cf. Kiparsky, 1973b, 95f.). Finnish has a defective segment, here transcribed //, which is realized identically to a following consonant if there is one, and is deleted elsewhere (as in 'go down' and 'go'). Phonologically, // may be analyzed as a bare consonantal root node.¹⁹

(24) /mene' pois/	menep pois	'go away'
/mene' talolle/	menet talolle	'go to the house'
/mene' kotiin/	menek kotiin	'go home'
/mene' vasemmalle/	menev vasemmalle	'go to the left'
/mene' alas/	mene alas	'go down'
/mene'/	mene	'go'

Scobbie (1993) notes that 'elsewhere' in the general rule may trivially be reinterpreted as the negation of the context of the particular rule, allowing such interactions to be remodeled declaratively, that is, without interaction. This is obviously impossible with rules whose contexts overlap, which requires extrinsic ordering.

2.7 Cyclicity

In theories of extrinsic ordering, each rule has only one chance to apply. In SPE, certain rules escape this restriction. The application of such **cyclic rules** is interleaved with processes that build successively larger morphological and syntactic constituents. Cyclic rule application begins with the innermost domain and proceeds outwards to the next most inclusive domain, and so on to the outermost. Certain nodes in morphological and syntactic structure, or the affixes that derive them, may be designated as cyclic. They trigger (re)application of cyclic rules (Brame, 1974). In English, for example, there is a basic division between suffixes that trigger reassignment of stress, and those that do not. Latinate suffixes typically trigger a new cycle, giving alternations like *vólume~volúminous~volumínosity*. Germanic suffixes, on the other hand, generally do not work this way, cf. *Hémulen~Hémulenish~Hémulenishness*.

Elision in Tangale is another example of a cyclic rule. Elision may create consonant clusters that are repaired by epenthesis (Kidda, 1993: 44). Representative examples are given in (25). The epenthetic vowel is a high back rounded vowel [u] or [ʊ] that harmonizes with the root for the feature [ATR].

- | | | | |
|------|------------|---------|-------------------|
| (25) | /tagza+nó/ | tagʊznó | ‘my sorghum’ |
| | /líprá+ží/ | lípʊrží | ‘your (f) needle’ |
| | /soqló+gó/ | soqʊlgó | ‘your (m) fish’ |

The derivation of *tagʊznó* ‘my sorghum’ is given in (26).

(26)	UR	/tagza+nó/
	Elision	tagznó
	Epenthesis	tagʊznó
	SR	tagʊznó

Evidence for the cyclic application of Elision comes from verbs, where the same personal suffixes that mark possession on nouns are used to mark the direct object. Direct object markers may be followed by additional suffixes marking tense, and which trigger new applications of Elision. The data in (27a) from Kidda (1993: 42, 44) illustrate cyclic Elision in verbs followed by an object and tense clitic, and (27b) verb followed by tense clitic and noun object.

- | | | | |
|------|---------------|----------|-------------------|
| (27) | | | |
| a. | /dóbé+nó+gó/ | dóbúngó | ‘called me’ |
| | /sugdé+ží+gó/ | sugdúžgó | ‘pierced you (f)’ |
| | /káyé+gó/ | káyúggó | ‘chased you (m)’ |

b. /pɔné+gó/	pɔngó	/pɔné+gó#ɲai/	pɔɲug ɲai	'knew (Ngai)
/aɲé+gó/	aɲgó	/aɲé+gó#ɲai/	aɲug ɲai	'paid (Ngai)
/timlé+gó/	timulgó	/timlé+gó#ɲai/	timlug ɲai	'hypnotized (Ngai)

The correct distribution of epenthetic vowels can be derived by assuming that Epenthesis is not a cyclic rule, applying instead after all the cyclic rules. The resulting derivation is shown in (28). Double brackets $\llbracket \ \rrbracket$ show the morphological bracketing. As the derivation moves outwards, the brackets of the previous cycle are erased.

(28) UR	$\llbracket \text{timlé go} \rrbracket$	$\llbracket \llbracket \text{timlé go} \rrbracket \text{ ɲai} \rrbracket$	$\llbracket \llbracket \text{sugdé ží} \rrbracket \text{ gó} \rrbracket$
First Cycle			
Elision	$\llbracket \text{timlgó} \rrbracket$	$\llbracket \text{timlgó} \rrbracket$	$\llbracket \text{sugdží} \rrbracket$
Second Cycle			
Elision	<i>n/a</i>	$\llbracket \text{timlg ɲai} \rrbracket$	$\llbracket \text{sugdžgó} \rrbracket$
Post-cyclic rules			
Epenthesis	$\llbracket \text{timulgó} \rrbracket$	$\llbracket \text{timlug ɲai} \rrbracket$	$\llbracket \text{sugdúžgó} \rrbracket$
SR	[timulgó]	[timlug ɲai]	[sugdúžgó]

Allowing Epenthesis to apply cyclically would predict the wrong result for 'hypnotized Ngai', as shown in (29).

(29) UR	$\llbracket \text{timlé go} \rrbracket$	$\llbracket \llbracket \text{timlé go} \rrbracket \text{ ɲai} \rrbracket$
First Cycle		
Elision	$\llbracket \text{timlgó} \rrbracket$	$\llbracket \text{timlgó} \rrbracket$
Epenthesis	$\llbracket \text{timulgó} \rrbracket$	$\llbracket \text{timulgó} \rrbracket$
Second Cycle		
Elision	<i>n/a</i>	$\llbracket \text{timulg ɲai} \rrbracket$
Epenthesis	<i>n/a</i>	$\llbracket \text{timulug ɲai} \rrbracket$
SR	[timulgó]	*[timulug ɲai]

Having to stipulate which rules are cyclic and which are not may seem like an undesirable enrichment of the theory. Because Elision specifically targets stem- or domain-final vowels preceding other phonological material, it is an example of a process that only applies in a morphologically **derived environment** (Kiparsky, 1976; Mascaró, 1976). Domain-finality does not have to be inscribed into the formulation of the rule. It falls out naturally from the assumption that it is a cyclic rule. Epenthesis does not have this restriction. Suppose we were to try and eliminate the distinction by claiming that all rules were cyclic. Then we might attempt to replace Elision with a purely

phonological Syncope rule that targets any medial light syllables irrespective of domain. This would serve to get rid of the unwanted epenthetic vowel in the counterfactual output *[timulug ηai] in (29), but it would also counterfactually predict the deletion of medial light vowels in morphologically non-derived environments. The existence of trisyllabic words such as *dánǰiki* 'blouse', *lanǰó* 'donkey' and *mbipíde* 'cobra', all with a medial light syllable intact, proves this assumption wrong: the medial light syllable has to be the result of the concatenation of morphemes in order for it to count as a target for deletion. In other words, we're back to our original Elision rule.

Cyclic Phonology was further elaborated in Lexical Phonology (Kiparsky, 1982b, 1984, 1985; Kaisse and Shaw, 1985; Mohannan, 1986), which attempted to formulate the general properties of cyclic and other levels, also known as **strata**. There is a broad consensus among practitioners of Lexical Phonology that there are three strata, two lexical and one post-lexical. The lexical strata are the stem level (cyclic) and the word level (post-cyclic). The counterbleeding interaction of Canadian Raising and Flapping requires that Raising apply in an earlier stratum than Flapping. The ascription to stratum should follow from the morpho-syntactic domain in which the process applies. Since Flapping applies across word-boundaries (as in *my mother bough[r] a parrot*), it can be ascribed to the post-lexical component. Raising, which doesn't look beyond the word for its context, is lexical. For our purposes, a more specific level ascription is unnecessary (although see Bermúdez-Otero (2004) for an argument that Raising should be assigned to the Stem Level). Another important innovation with respect to Cyclic Phonology was that Lexical Phonology also manifested a concern with the properties of levels as levels, as opposed to mere steps in a derivation. Lexical rules, for example, are restricted to outputting structures that are also possible in underlying representation: they must be **structure preserving**. Post-lexical rules may produce allophones, even gradient outputs. Lexical rules may have exceptions, but post-lexical rules apply without exception. Finally, post-lexical rules are blind to domain, applying also across word boundaries. Since Canadian Raising is lexical, it must be structure preserving, that is, it outputs segments that are possible underlying forms. As we shall see in Section 3, this prediction is borne out by the existence of words like *tiger*, idiosyncratically with [AY], which lack the Raising context.

Stratal OT is an adaptation of the central ideas of Lexical Phonology in OT (Kiparsky, 2000, 2002; Bermúdez-Otero, to appear). Each stratum corresponds to an OT grammar, each a total ranking of the universal set of constraints. An important difference between rule-based Lexical Phonology and Stratal OT, however, is the predictions each makes with respect to within-cycle phonological opacity. In rule-based LP, each stratum consists of an extrinsically ordered battery of rules, allowing the capture of within-cycle counterfeeding and counterbleeding relationships. In Stratal OT, however, each stratum is transparent.

The Tableau 6.5 illustrate Stratal OT applied to the interaction of Canadian Raising and Flapping. The strata differ in the relative rankings of IDENT[son] and *VTV. On the Stem level, Flapping does not take place because IDENT[son] outranks *VTV. This ranking is reversed on the Post-lexical level, which produces Flapping. The overapplication effect is derived by ranking IDENT[low] high on the Post-lexical level, which forces identity of vowel height with the output of the previous level.

Tableau 6.5

a. *Stem Level*

/rāyt + ŋ/	RAISING	IDENT[son]	IDENT[low]	*VTV	*Λ
a. rāytŋ	*!			*	
b. r rAyŋ			*	*	*
c. rāyŋŋ		*!			
d. rAyŋŋ		*!	*		*

b. *Post-lexical Level*

rAyŋ	RAISING	*VTV	IDENT[low]	IDENT[son]	*Λ
a. rāytŋ		*!	*		
b. rAyŋ		*!			*
c. rāyŋŋ			*!	*	*
d. r rAyŋŋ				*	*

Parallel OT has developed other approaches to cyclic effects, such as Transderivational Correspondence Theory (Benua, 1997). Actually occurring words, like *write* and *writing*, enter into correspondence relations with each other, permitting the formulation of Output-Output Faithfulness constraints that require identity (in some dimension) of bases and their derivatives.

3. Alternatives to Opaque Interaction

In the previous section we compared and contrasted theories for which the ordering of processes or rules has to be specified in the grammar of a particular

language (extrinsic ordering) and theories in which the sequence of application is made to follow from principles of Universal Grammar (intrinsic ordering). From the point of view of explaining language acquisition, theories where as much as possible derives from universal principles are to be preferred, for the reason that they impose restrictions on the space of hypotheses that the learner can posit in the course of constructing the grammar of their native language. The kind of theory that is most interesting from the point of view of explaining language acquisition, however, is not necessarily best from the perspective of attaining maximal empirical coverage. This tension between the demands of explanation and description are evident in the debates on extrinsic and intrinsic ordering. Extrinsic ordering is explanatorily costly because it enlarges the learner's space of possible solutions, but perhaps necessary, if it turns out that nothing else can account for the attested range of process or rule interactions. Opponents of extrinsic ordering have appealed to four kinds of alternative account, listed in (30); which, if any, of these strategies is appropriate can only be determined on a case by case basis, however. Specific types of opacity do not have any single type of reinterpretation.

- (30) a. phonological reinterpretation (representations; other rules)
- b. phonetics
- c. allomorphy/suppletion
- d. analogy/lexicon

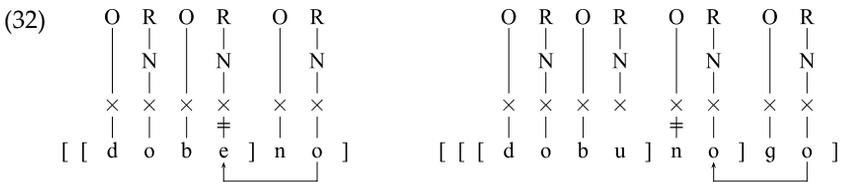
Given an alternation A-B, and a hypothesized rule $A \rightarrow B/C_D$, one solution is to retain the alternation in the phonology but give it a different interpretation. We can identify two subtypes of this strategy. The first is to try and replace the opaque rule with a transparent one. Sometimes there may be an alternative surface-apparent condition that we have missed which allows us to write a different, fully transparent, rule, say, $A \rightarrow B/E_F$. The second is to reinterpret the alternation in representational terms so that, despite appearances, the 'opaque' structure does not satisfy the structural description of the rule. The three remaining strategies all involve reallocating the alternation to a different component. For example, it may not be the case that B should figure in the discrete surface representation at all, but be allowed to emerge from the way this representation is mapped onto continuous parametric representations in the phonetic component (e.g. Browman and Goldstein, 1986). The last two strategies may be traced to the work of Mikołaj Kruszewski and the Kazan School (Kruszewski, 1881; Baudouin de Courtenay, 1972 [1895]; Klausenburger, 1978; Anderson, 1985). Closer investigation may reveal that B is not a phonological alternant of A (approximating Kruszewski's 'divergent'), but that A and B involve suppletive allomorphy (cf. Kruszewski's 'correlative'). We will exemplify each of these in turn using examples adduced earlier in the chapter.

Staying within the phonological ambit for the time being, consider a potential transparent reinterpretation of an opaque rule. Poliquin’s argument for Duke-of-York derivations in Canadian French depends on understanding laxing in non-final high vowels as the result of a Vowel Harmony rule.²⁰ As interpreted by Poliquin (p. 179), a high vowel in an open syllable becomes [-ATR] when separated by zero or more syllables from a word-final closed syllable containing a [+high, -ATR] vowel. The rule formulated by Poliquin is given in (31).

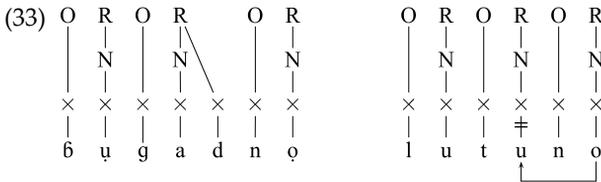
$$(31) [+high] \rightarrow [-ATR]/___ \sigma . (\sigma_0) . [+high, -ATR]C(C) \sigma \# \quad (\text{optional})$$

Interpreted this way, the rule of Pre-voiced Fricative Tensing inevitably destroys part of the context for the application of Vowel Harmony, and the result is apparent counterbleeding. Another interpretation of the facts, though, is that the required rule is not a Vowel Harmony rule at all, but an entirely transparent rule that changes a non-final high-vowel to [-ATR] preceding any high vowel in the last syllable, irrespective of its specification for [ATR].

The handling of vowel-zero alternations in Government Phonology may serve as an illustration of how representations may be used to reanalyze putative opacity effects. Consider once again the facts of Elision and Epenthesis in Tangale. Building on Nikiema (1989a and b), Charette (1991: 106–112) provides an analysis of Tangale Elision in which the vowel is delinked from the nucleus but the position itself remains – as required by the principle of Government Phonology that every representation begin with an O(nset) and end in a R(hyme). For *dobno* ‘call me’ and *dobungo* ‘called me’, Charette proposes the structures in (32). When the empty nucleus is properly governed by a following nucleus with melodic content, it has no phonetic interpretation. When it is not properly governed, the empty nucleus is interpreted as [u].



empty nucleus inhibits spreading of the relevant feature. This blocking effect is illustrated in (33). Solutions of this kind, then, redefine the representation in such a way that the structural description of the would-be opaque rule cannot be met.



McCarthy's (2006) response to putative rule sandwiching in Mohawk illustrates how certain phenomena might be reassigned to phonetic interpretation. The way in which the phasing of **articulatory gestures** may give rise to apparent deletions and insertions is a fruitful area of research activity (Browman and Goldstein, 1986). Overlapping gestures may partially or completely mask certain segments, even though represented by the appropriate gesture. In a phrase like *ten pin*, the labial closure gesture will in casual speech overlap with the preceding nasal, resulting in something that sounds like [tem pin], although /n/ retains an apical gesture in articulation. In such cases, it would be inappropriate to speak of the segment as having been deleted. Similarly, incomplete overlap, or open transitions between gestures may result in the appearance of excrescent vowels (cf. Levin, 1987). Such vowels are not epenthetic in a phonological sense, since they are not present in the surface representation. Hall (2003) argues that excrescence is more likely to occur in transitions to voiced sounds. McCarthy observes that what we called Biliteral Epenthesis above, the second consonant must be one of the resonants /n r w/. He argues that Mohawk has only one epenthesis rule, Triliteral Epenthesis, which interacts transparently with stress. The *e* that is invisible to stress is simply an excrescent vowel with no phonological representation.

This brings us to the morpholexical solution. Several attempts have been made to bring it to bear from within the OT camp, including Sanders (2003), Mielke et al. (2003), and Green (2004, 2007). As we saw in Section 1, the opaque interaction of Canadian Raising and Flapping played a major role in the emergence of the generative paradigm in phonology. Attempts to eliminate the opacity date right back to Joos' seminal paper (Joos, 1942). Kaye (2008) presents a recent argument against opaque Canadian Raising, which he argues is an artefact of the assumption that the diphthongs in *ride* and *write* are underlyingly the same. He notes first of all that, considered apart from any high off-glides, the vowels [ʌ]

and [a] are actually contrastive in Canadian English, as evidenced by minimal pairs such as [kʌ'b] *cut* versus [kɑ'b] *cut* and [kʌt] *cut* versus [kɑt] *cut*. Given the independent existence within the system of both of these vowels, Kaye suggests (p. 17) '[t]he obvious analysis is that C[anadian] E[nglish] "right" and "out" have rʌyt and ʌwt as their lexical representations', that is, the initial vowels of [ʌy] and [ʌw] now belong to the phoneme /ʌ/, not /ɑ/.²¹ Kaye concludes that deriving [ʌy] from [ɑy] would introduce an unnecessary disparity between the underlying and surface levels: the qualitative difference therefore cannot be allophonic. If the traditional phonological account is correct, however, we would expect to find [ʌy] and [ʌw] limited to the environment preceding a voiceless obstruent. What we find instead, at least for [ʌy], is that it can occur lexically in non-raising environments. Kaye reports that some speakers use this quality in *tiger*. K. Currie Hall (2005), working within a stochastic, exemplar-based approach, also reports a breakdown, especially noticeable in the case of polysyllabic words, in the predictability of the distribution of [ɑy] and [ʌy]. She demonstrates that lexical neighbours (words that share similar or identical substrings) may in fact exert an analogical influence on the choice of the diphthong in a way that runs counter to the putative phonological generalizations. Thus, *gigantic* and *angina* unexpectedly have [dʒʌy], despite the absence of a voiceless obstruent in the relevant syllable, and *Siberia* and *psychology* have [sʌy]. Completing Kaye's point, she shows that [ɑy] turns up in syllables closed by a voiceless obstruent, such as *like* and *life*. In her 2005 paper, Currie Hall suggests factoring lexical neighbourhoods into the description of the phonological rule, but in her 2009 dissertation she essentially opts for relexicalization:

Perhaps having contrast in 4% of environments (before [r] in words like *writing/riding*) has opened the door for new generalizations to emerge. For example, language users could generalize that [ʌi] is possible before voiced segments and extend that to other words like *gigantic*. The prediction then is that [ɑi] and [ʌi] could continue along the continuum and end up being entirely unpredictably distributed: Fully contrastive. (Currie Hall, 2009: 124)

This perspective opens up interesting possibilities for understanding structure preservation phenomena. Recall from the previous discussion that opaque rules are necessarily lexical and, being lexical, they can only output contrastive segments of the language. When allophones are opacified, therefore, they are necessarily reanalyzed as phonemes. However, structure preservation is only a stipulated property of lexical rules. The ultimate reason for the breakdown observed by K. Currie Hall may be that children acquire their knowledge of phonological distributions before undertaking any morphological analysis of the input, essentially applying structuralist discovery procedures to construct the inventory of contrastive segments in their language. Some words thus come

over time to be lexicalized with the 'wrong' allophone. One possible mechanism underlying errors of this type may be analogy with forms where the allophonic rule has overapplied. Alternatively, in the absence of a 100 per cent reliable generalization learners may be unsure about which sound category to assign a given word, especially one that they do not encounter frequently. When the child later acquires the morphology and opaque rules along with it, it is too late to undo relexicalization. The possibility of relexicalization raises the interesting question whether children do in fact acquire opaque rules at all or whether the generalizations are simply part of the lexicon. See however McCarthy (2006) and Ettliger (2009) for arguments for the psychological reality and learnability of opaque rules.

What is allophonic according to Kaye is the variation in the length of the vowel: all stress bearing vowels in monosyllables lengthen in what Kaye calls the 'non-fortis environment', which is to say at the end of a word, optionally preceding a consonant not drawn from the set of voiceless obstruents. Thus, [bæt] *bat* versus [bæːd] *bad*, [kāt] *cot* versus [kɑːd] *cod*. This kind of variation is entirely general in English. Since Kaye's Non-fortis Lengthening and Flapping are mutually non-affecting, however, their relative ordering is simply not crucial.²²

4. Conclusions

A derivation is a mapping between an underlying and a surface level, each of which is made up of the same basic type of representational elements. Derivations also entail the application of processes that potentially alter underlying structure in destructive ways. As the field has developed, the relative emphasis on levels and derivations has shifted. At certain times, most obviously at the inception of the generative programme in phonology, a strong emphasis on derivations, processes and rules has tended to divert attention away from an interest in the properties and functions of levels and their interrelation. Nevertheless, a concern with levels has repeatedly reasserted itself, and changes in the conception of levels have brought about corresponding changes in the understanding of derivations. The fundamental question remains whether linguistically significant generalizations should be stated as constraints on one or more levels, or as rules deriving one level from another. It is interesting in this connection to compare rule-based phonology and OT in their earliest and most radical versions with the more moderate positions most practitioners of each approach have come to adopt in the light of experience with real data. Most versions of rule-based phonology came to incorporate both restrictions on underlying representations (morpheme structure conditions), surface representations (output constraints such as the OCP), and outputs of lexical strata

(Structure Preservation). OT has in its turn had to find ways of mimicking the extrinsic ordering of processes to deal with the opacity problem. Perhaps it is too optimistic to hope that a consensus on these fundamental issues will emerge any time soon. But between these two ways of understanding the grammar is a creative tension that will surely spur us on to novel insights about the nature of language for decades to come.

5. Acknowledgements

I would like to thank Martin Krämer, and two anonymous reviewers for their invaluable feedback.

6. Notes

1. Several early implementations of generative grammar do in fact posit uninterpretable symbols in underlying or deep structure. Such symbols had to be deleted during the course of the derivation. Indeed, the derivation was partly driven by the requirement that such uninterpretable symbols be absent at the surface. In Chomsky (1965), for example, sentences that had not (yet) been incorporated into a higher matrix clause were bounded by the juncture symbol #. Every occurrence of # in the deep structure *John saw the man # the man kicked the cat #* must be eliminated by applying a relative clause transformation to derive *John saw the man who kicked the cat*. Similarly, in chapter 9 of SPE, distinctive features could only have values of + or – at the surface, but underlyingly they might also be specified as ‘marked’ *m* or ‘unmarked’ *u*. In the course of the derivation, these values were replaced by + and – through what were known as marking conventions. The notion of unmarkedness continued to develop in Underspecification Theory (see Steriade, 1995b, for an overview) and representational theories allowing only unary features, such as Government Phonology.
2. As Goldsmith notes, this conception had a couple of ready metaphors to hand, in the derivation from proto-forms in historical linguistics and in production systems used in artificial intelligence.
3. Other terms commonly in use in the literature for ‘focus’ and ‘environment’ include, for ‘focus’, **target**, and for ‘environment’, **context**, **determinant**, or **trigger**.
4. The origins of this conception are due to Bloomfield (1939), who cites the *Aṣṭadhyāyī* of Pāṇini as his chief inspiration, and Wells (1949). See Goldsmith (2008) for discussion of pre-generative antecedents of this concept.
5. These terms are ultimately due to Kiparsky (1968).
6. McCarthy (2007a: 101) proposes the term **designated level** to refer to those representations ‘with special restrictions on their content or unique roles to play’. In addition to the underlying and surface levels, the output of a cycle also constitutes such a designated level. See Section 2.7.
7. A vacuous Duke-of-York derivation would skip \mathcal{Q} , and \mathcal{R} would destroy \mathcal{S} as soon as \mathcal{P} created it. While possible formally, the learner would lack the crucial evidence to posit such an interaction.
8. As the Austronesian scholar Ken Rehg comments (Rehg, 2001: 218), though, ‘one of the imperatives of generative phonology – that allomorphy must be minimized – is

sometimes at odds with the data, typically in very subtle ways'. Suppletion is easy to diagnose when the variants in question have little in common phonetically, but suppletion cannot be ruled out even in the case of phonetically similar variants. There is a widespread assumption that phonetically similar allomorphs must derive from the same underlying form, and this very assumption will tend to favour analyses with abstract underlying representations and rule interaction. Researchers working within Government Phonology have also pointed out that adhering to this assumption in all cases may run into empirical problems, not to mention the theoretical issues raised by abstractness and extrinsic ordering.

9. Schane cites a dissertation by Flora (1974).
10. As a terminological caveat, note that a pair of rules may interact transparently, but either of the rules may be opaque because it is opacified by some third rule.
11. These terms are originally due to Wilbur (1973).
12. A further logical type is what Koutsoudas (1976: 4) terms **arbitrary ordering**, which resembles random ordering except that the rules may not apply more than once.
13. Baković (2007, to appear) argues that this does not represent a genuine case of bleeding. Baković raises concerns about implicit duplication in the traditional way of understanding the environment of epenthesis as applying between two consonants that potentially differ only in their value for the feature [voice]. He writes, 'Epenthesis must arbitrarily ignore the difference in voicing between the stem-final /t/ and the suffix /d/ – precisely the difference that would be neutralized by Assimilation were it to apply. A more explanatorily adequate analysis would eliminate this redundancy, making *strict identity* a requirement on Epenthesis [. . .] and relying on Assimilation to provide the necessary context [. . .]'. Baković uses this interpretation of epenthesis to argue for a type of opacity he dubs **cross-derivational feeding**. Showing how this works would take us too far afield, but suffice it to mention that his concern is to show that OT is able to handle some types of opacity that rule-based theories cannot. *Contra* Baković, I can't find any reason for thinking that allowing epenthesis to ignore [voice] is arbitrary in any way. That homorganic similarity rather than complete segmental identity is the correct interpretation is also suggested by the fact that epenthesis in English applies to breakup clusters of sibilants that potentially differ not only in their specification for [voice] but also in their value for the feature [anterior], for example, [bʊʃɪz] *bushes*. In the case of [anterior], the issue of duplication does not arise, since there is no phonological assimilation rule in English involving this feature.
14. Myers (1991) argues for an approach with a division of labour between extrinsically ordered and persistent rules, stating (p. 339) 'the set of persistent rules specifies the inventory of phonological elements (segments, syllables and feet) in a language, whereas the ordered rules specify the distribution of those elements'.
15. McCarthy (2007b) originally proposed the name 'Persistent OT'.
16. A partial solution to this problem is offered by Ettliger's 2009 Diagonal Correspondence Theory (DCT). DCT builds on proposals in the literature that segments within the same harmonic span may be 'syntagmatic correspondents'. In this way, the focus of the autosegmental spreading of some feature may be the 'diagonal correspondent' of that feature's underlying determinant. The underlying determinant of a spreading process may thus have diagonal correspondents in the surface representation whether or not it has 'perpendicular' correspondents (it has been deleted), and whether or not its perpendicular correspondent has identical values for the harmonizing feature. This allows Ettliger to describe opaque spreading rules. It is not clear how it may be extended to other types of rule, however.
17. There will always be other plausible candidates that would need to be excluded by bringing on additional constraints in a fully worked out account. In the interests of

space, though, we abstract away from this here and simply offer the barebones of an analysis for each case.

18. Other intrinsic principles include Archangeli's (1984: 51) Redundancy Rule Ordering Constraint, which forces all redundancy rules to apply as late in the derivation as possible.
19. Kiparsky identifies the segment as /k/, but this choice appears to be synchronically arbitrary. The only evidence available to the child is that the segment is an unspecified consonant of some kind.
20. I am extremely grateful to an anonymous reviewer for bringing this important point to my attention.
21. The relexicalization argument is rendered plausible by the existence of other clear cases of contextual relexicalization in the history of English. In certain parts of the English-speaking world, including the southeast of England, southern New England and eastern Virginia, [æ] was lengthened to [æ:] and subsequently retracted to [ɑ:] before the fricatives [f θ s] and coronal nasal+obstruent clusters [nt nd ns nʃ]. Despite the importance of context, the sound change was clearly not allophonic, as evidenced by the existence of lexical items that failed to undergo the reassignment, such as *asp*, *mass* and *pant*. The replacement of /æ/ with /ɑ:/ in many words is a clear case of **lexical diffusion** (Wang, 1968).
22. Idsardi (to appear) nevertheless provides evidence in a recent account that Canadian Raising is a productive rule that results in actual alternations, for example, [aɪ] *i* versus [ɪθ] *i*-th and [waɪ] *y* versus [wɪθ] *y*-th. He also mentions that his own variety displays raising before certain clitic prepositions, for example, [laɪ] *lie* versus [lɪɪθ mi] *lie to me*.

7 Constraint-Based Phonology

Christian Uffmann

Chapter Overview

1. Introduction	174
2. The Rise of Constraints in Phonological Theory	176
3. The Emergence of Constraint-Based Frameworks	183
4. Optimality Theory	188
5. Discussion	199
6. Notes	201

1. Introduction

Every phonological process can be viewed from two perspectives, the perspective of the input to the process or the perspective of the output. Traditional rule-based phonology takes the input-based perspective: rules are statements about which changes underlying representations undergo in their transformation to surface representations. The role of the input is crucial because it determines whether a rule will apply or not (the structural description is a description of the input). Conversely, an output-based perspective looks at the kinds of structures a process generates on the surface – or avoids. In phonological theory, such conditions or restrictions on possible outputs are captured by **constraints**.

To make this distinction clearer, consider an example of a well known and cross-linguistically frequent type of process, final devoicing, where underlyingly voiced word-final or coda obstruents become voiceless. The following data (1) from German illustrate this process: underlying /d,b,z/ surface as [t,p,s] in word-final position.¹ In (b), non-alternating forms are shown, which demonstrate that the process in question is not one of intervocalic voicing.

(1) Final devoicing in German

	<i>singular</i>	<i>plural</i>	<i>gloss</i>	<i>UR</i>
(a)	[kʌ:t]	[kʌ:ɖə]	'wheel'	/kʌ:d/
	[di:p]	[di:bə]	'thief'	/di:b/
	[gʌ:s]	[gʌ:zə]	'grass'	/gʌ:z/
(b)	[kʌ:t]	[kʌ:tə]	'council'	/kʌ:t/
	[fas]	[fesə]	'vat'	/fas/

The traditional generative perspective accounts for this alternation by stating a maximally general rule that describes the change from underlying to surface form, as in (2).

$$(2) \left[\begin{array}{c} -\text{son} \\ +\text{voice} \end{array} \right] \rightarrow [-\text{voice}] / ___\# \#$$

This rule states that any voiced obstruent will become voiceless in word-final position. From a constraint-based perspective, there is a different generalization to be made: that German prohibits voiced obstruents in word-final position. This generalization can be stated as a negative constraint, that is, as a statement that acts as a filter on possible outputs, as in (3):

$$(3) * \left[\begin{array}{c} -\text{son} \\ +\text{voice} \end{array} \right] \# \#$$

A juxtaposition of (2) and (3) shows that rules and constraints have different properties. The rule in (2) accounts for the process in (1) in describing the change that forms undergo while the constraint in (3) in a sense motivates the change. Why does the rule in (2) exist? Because there is a constraint against final voiced obstruents. This in turn raises the question of why the constraint exists and why final devoicing is a cross-linguistically common process while its converse, final voicing, is not (it is robustly unattested, *pace* Blevins, 2004). Do constraints embody universal phonological and phonetic principles?

Before we move on, notice another point: It is possible to describe the phonological alternations of a language just by a set of rules. The constraints may add some explanatory adequacy, but they are not strictly needed. The converse does not seem to be the case: the constraint in (3) does not fully account for the process in (1). It may motivate it but it does not say anything about the kind of process employed to repair the illegal structure. (3) could as well be satisfied by deleting the offending segment, or by adding a final vowel which would render the voiced obstruent no longer word-final. The fact that final devoicing is found as a repair, rather than any other process which would ensure that the

final segment is not a voiced obstruent, is outside the scope of the constraint. It is no wonder, then, that early generative theory only used rules to account for all phonological alternations and generalizations. They are sufficient to formalize all processes, and consequently constraints had no formal status in phonological theory.

In this chapter, we will look at how constraints entered phonological theorizing and how their role in phonology grew more and more important, leading to constraint-based models of phonology which dispose of rules altogether, the best known and most successful of these models being Optimality Theory. The next section will briefly review the development of constraints in phonological theory, followed by an overview of constraint-based approaches emerging in the 1980s. This is followed by a more detailed introduction to Optimality Theory, which is the arguably most influential constraint-based model of grammar and which has been the dominant paradigm in phonological theory since the mid 1990s.

2. The Rise of Constraints in Phonological Theory

Constraints entered the formal phonological system – and started playing a central role in it – essentially via two different paths, both of which will be outlined briefly in this section. On the one hand, the insight was growing that constraints can be the locus of linguistically significant generalizations; theories that reject constraints as part of the formal system could express these generalizations only indirectly, for example, via the interaction (or *conspiracy*) of rules. On the other hand, the rule formalism itself was increasingly seen as too powerful, in its ability to formally describe unattested types of languages. So again, constraints were brought in to curtail the expressive power of rules, to yield a closer match between the formal possibilities of the theory and empirically observed variation.

2.1 The Explanatory Power of Constraints

Processes or alternations are not the only place where phonological generalizations hold. There are also phonotactic generalizations; for example, many languages allow only open syllables. In English, any triconsonantal cluster word-initially is subject to heavy phonotactic restrictions: the first element must be /s/, the second element a voiceless stop /p,t,k/, the third element a liquid /l,r/, with the additional restriction that /tl/ is an impossible cluster.² This reduces the number of possible trilateral clusters to /spr, spl, str, skr, skl/ (and /skl/ is marginal, occurring only in loanwords).

Such a generalization must be captured somewhere in the phonological system. A hypothetical initial cluster /kls/ is not only robustly unattested in English; every native speaker intuitively rejects such a cluster as non-English. We therefore need to exclude initial /kls/ from the set of possible underlying forms. In earliest generative theory, this was entirely done by context-sensitive redundancy rules, so-called morpheme structure rules (Halle, 1959). These come in two shapes. First, to use the terminology in Harms (1968), there are *blank-filling rules*, context-free redundancy rules, as known from underspecification theory: features that are non-contrastive for a segment are not specified in the lexicon. Hence, /t/ is not specified for [+anterior, +distributed] because it is the only [+coronal] stop in the inventory. The second type of rules are *sequential constraint rules*. They further underspecify the feature matrix of a segment by taking phonotactic information into account. The example in (4) shows an underspecified matrix for the cluster /str/. Only the feature specifications given here are necessary to unambiguously distinguish this cluster from all other triconsonantal clusters. The task of the phonologist, then, is to establish the rules that can underspecify the feature matrix in order to account for all phonotactic restrictions.

(4) An underspecified /str/ cluster

$$[+cons] \left[\begin{array}{c} +cons \\ +coronal \end{array} \right] [+cons]$$

In detail: It is sufficient to specify the first segment as just [+consonantal] because there is only one consonant that can appear in this position (all triconsonantal clusters begin with /s/). /t/ only needs to be specified as [+coronal] because only voiceless stops can occur in second position of the cluster. Finally, there is no need to specify the third consonant; the only trilateral cluster beginning with /st/ is /str/. To take an example, a redundancy rule for determining that /s/ is the first consonant in a trilateral cluster is found in (5).

(5) A redundancy rule for /s/ in /sCC/ clusters

$$[+cons] \rightarrow \left[\begin{array}{c} -sonorant \\ +continuant \\ +coronal \\ +anterior \\ +distributed \\ -voice \end{array} \right] / _ _ _ \# \# [+cons] [+cons]$$

These rules, as Karttunen (1969) points out, have to be bidirectional. Not only are they used to fill in predictable features, they also have to be able to remove

redundant features from a hypothetical overspecified input, to avoid introducing new contrasts into the lexicon. Underspecification is thus not just optional or driven by an abstract idea of representational economy; it is necessary to rule out unattested segments or segment combinations in a language.

The rule-based approach to phonotactic restrictions in the lexicon is thus far from straightforward. There is a battery of highly specific rules that underspecify lexical entries from which generalizations emerge only indirectly. Nowhere is it stated directly in the grammar that a triconsonantal initial cluster must start with /s/; it is an epiphenomenon of a number of context-sensitive redundancy rules. For this reason the proposal in Stanley (1967), that phonotactic restrictions be formalized not by rules but rather by morpheme structure constraints (MSCs) (or *conditions* in his original formulation) was quickly adopted. One of his main concerns was the ‘improper’ use of blanks (features with a zero value) in morpheme structure rules, which meant that binary features would in practice be ternary, with unclear consequences. In addition, he shows how morpheme structure rules can impose empirically unattested types of phonotactic constraints on words. Stanley proposed three types of MSCs, positive, negative and if-then, which became the standard way of formalizing constraints. Positive conditions state which structures are allowed. They can be seen as a filter through which only those structures pass that match its description. For the triconsonantal cluster restriction in English, for example, a positive constraint as in (6) can be formulated, which states that the first segment must be /s/, the second a voiceless stop, the third a liquid.

(6) Positive MSC on triconsonantal initial clusters in English

$$\# \# /s/ \left[\begin{array}{l} +\text{consonantal} \\ -\text{sonorant} \\ -\text{continuant} \\ -\text{voice} \end{array} \right] \left[\begin{array}{l} +\text{consonantal} \\ +\text{sonorant} \\ +\text{continuant} \end{array} \right]$$

A negative condition is like the one we saw in (3) for final devoicing (although (3) is not a morpheme structure condition since it holds for surface forms): such conditions state explicitly what is not allowed, usually in the format of an asterisked structure. They thus work as a filter which blocks every structure matching its description. Finally, if-then conditions come in a shape reminiscent of rules (and Stanley points out that they are convertible to rules). In our example of triconsonantal clusters in English, one could state that if there is a cluster of three [+cons] segments, then the first one must be /s/ (and see (7)).

While Stanley’s proposal was quickly accepted (see, for example, Chomsky and Halle, 1968), the question emerged whether constraints play a larger role in phonology, besides restricting the shape of lexical entries. This possibility was

explicitly denied in Postal (1968), who pointed out that all surface forms can be derived via a combination of MSCs and phonological rules. Therefore, additional constraints, such as constraints on surface forms, were empirically unnecessary. Nevertheless, this view was challenged by adducing arguments other than those of empirical necessity, arguments about where in the phonological system relevant generalizations are located.

An important step in this direction was the discussion of Yawelmani Yokuts in Kisseberth (1970a). Kisseberth points out that there are a number of insertion and deletion rules in this language, which formally come in quite different shapes but are functionally intimately related in that they all create licit syllable structures, by reducing clusters and avoiding hiatus. He uses the term *conspiracy* to describe such a situation in which different rules create similar output structures. Kisseberth (1970a: 306) also affirms that the truly interesting generalization about Yawelmani phonology lies in this conspiracy of functionally related rules, not in the statement of a range of formally different rules, and concludes that individual rules in a grammar may not be arbitrary but 'motivate to some extent the existence of another rule', and that the exploration of such connections between rules should be a major goal of phonological theorizing. He does not, however, introduce the notion of output constraints as a formal mechanism to express functional relatedness. This notion was independently proposed by Shibatani (1973) and Sommerstein (1974) a few years later.

Shibatani and Sommerstein both argue for the existence of surface structure constraints (SSCs) in phonology, in addition to the MSCs proposed by Stanley, with similar arguments. For both, the main motivation for SSCs (or SPCs, surface phonetic constraints, in Shibatani's terms) is psychological: If a theory of grammar is to model the knowledge of speakers, it should contain SSCs. They argue that speakers' knowledge of allowed and disallowed structures is based on surface properties of strings alone, not MSCs. The knowledge of German speakers that there are no word-final voiced obstruents is surface-based; underlyingly, German does allow word-final voiced obstruents. Similarly, the oft-cited point that every native speaker of English knows that *blick* is a possible, though non-existing, word while **bnick* isn't, can only be based on SSCs because tautosyllabic obstruent-nasal clusters are only disallowed at the surface level, not underlying, as demonstrated by alternations such as *sign-signal*, *signify* or *gnosis-agnostic*. These show that the underlying forms do contain such clusters, which can therefore not be banned by MSCs. In addition, SSCs are the locus at which generalizations can be made regarding the type of rule conspiracy discussed in Kisseberth (1970a), by prohibiting hiatus or excessive consonant clustering.

Once one allows constraints into the formal system of grammar to complement rules, however, the question arises as to how the two interact, especially

when, as stated in the initial paragraphs of this chapter, constraints may be said to motivate rules. Both Shibatani and Sommerstein offer proposals how the use of constraints may simplify the rule component. Shibatani uses if-then constraints to formalize surface structure restrictions, pointing out that they are easily translated into a rule; such pairings of a constraint with a rule of identical form he calls an alternation constraint, or A/SPC. To take an example, consider the following if-then constraint to express final devoicing in German:

- (7) Constraint triggering final devoicing
IF : [-sonorant] ##
↓
THEN : [-voice]

This constraint translates straightforwardly into the rule in (2). Sommerstein (1974, 1977) goes a step further: He suggests that the conditioning environment be removed from the rule formalism altogether so that rules state only the structural change (e.g. [+voice] → [-voice] for final devoicing), with the added condition that a rule ‘does not apply unless its application will remove or alleviate a violation or violations of C [the constraint]’ (Sommerstein 1974: 75). This is an important reconceptualization of the notion of phonological rule as a repair strategy to remove (or to lessen) a constraint violation. It also challenges the view that phonological derivations consist of a set of ordered rules. The rules no longer are ordered but are available whenever a constraint violation calls for a repair. Sommerstein’s proposal thus foreshadowed models of phonology in the late 1980s and early 1990s to which we will turn presently. Before that, a second line of reasoning has to be discussed briefly that also contributed to the development of constraint-based phonology, the question of universals in phonology and the problem of constraining the rule format.

2.2 Constraining Rules

An important motivation for including constraints within the formalism of generative phonology was the power of the rule formalism, which many phonologists perceived to be excessive. Chomsky and Halle (1968) are explicit about the problem in the final chapter of the *Sound Pattern of English*. They point out that if they replaced all [+F] specifications in their account of English with [-F] specifications, and vice versa, this would still be a formally possible grammar, although a highly unlikely, if not impossible one. They conclude that ‘to the extent that this is true, we have failed to formulate the principles of linguistic theory, or universal grammar, in a satisfactory manner’ (Chomsky and Halle, 1968: 400), especially since an overly formal theory ignores the

'intrinsic', or substantive content of distinctive features. Chomsky and Halle thus raise a question which would come to play a pivotal role in phonological theory: how are the notions of markedness and naturalness integrated into a theory of phonology, given that they are an important source of linguistically relevant generalizations?

The answer they suggest is rule-based again. They propose that features cannot only have + or - values but can also be valued *u* (unmarked) or *m* (marked) and suggest a set of universal 'marking conventions' that specify what the context-dependent marked or unmarked values of a feature are. For voicing, for example (and we depart from their marking conventions here for ease of exposition), one could suggest a marking convention which states that the unmarked value for [voice] is [+voice] in sonorants (sonorants are typically voiced) and [-voice] in obstruents (voicelessness being unmarked in obstruents):

(8) A marking convention for voicing

$$[u \text{ voice}] \rightarrow \left\{ \begin{array}{l} [-\text{voice}] / \left[\begin{array}{c} \text{---} \\ -\text{son} \end{array} \right] \\ [+ \text{voice}] / \left[\begin{array}{c} \text{---} \\ +\text{son} \end{array} \right] \end{array} \right\}$$

Chomsky and Halle's proposal was not widely adopted. Especially problematic seems to be the introduction of *u* and *m* feature values to a theory of supposedly *binary* features, which now could take five values: +, -, 0 and *u/m*. Nevertheless, it put the important question on the agenda of how a theory of phonology interfaces with a theory of markedness or naturalness. This led to the emergence of Natural Phonology (Stampe, 1972a, see also Dinnsen, 1979) and Natural Generative Phonology (e.g. Hooper, 1976; Vennemann, 1974a) with their focus on sets of natural rules. Two ideas underlie these models: that the set of (universally) available operations is limited to a set of phonetically natural operations, and that all processes should be surface-true, that is, motivated from the output, which entails a rejection of extrinsically ordered rules and the opacity effects this ordering can cause. While these frameworks were criticized because of empirical evidence for 'unnatural' rules (Anderson, 1981) and for rule ordering effects (e.g. Kiparsky, 1968), they nevertheless firmly established the issue of markedness in phonological theory. As such they were an important influence on the development of Optimality Theory, and their challenge of the idea of ordered rules has remained attractive, also to those theorists less interested in naturalness. (Some examples of relevant approaches will be discussed in Section 3.)

A more general line of research thus emerged with the aim of curtailing the expressive power of rules. Here, constraints would again prove important, first,

as conditions for constraining the rule formalism; secondly, by raising the aforementioned question of how cross-linguistic restrictions – or tendencies – could be incorporated into the formalism.

Apart from constraining the rule formalism itself, for example, by curtailing the use of alpha notation, constraints were formulated as conditions on or principles of rule application, especially regarding rule blocking. A famous example is the Elsewhere Condition (Kiparsky, 1973b) that states under which conditions rules may be ordered disjunctively, as either–or rules (namely, if and only if the conditioning environment of one rule is a proper subset of the conditioning environment of another rule; see also Section 4.4). Another principle is that of Structure Preservation (Kiparsky, 1985), which captures the insight that a rule may be blocked from applying if the output of the rule is not in the inventory of permissible structures of the language. To take an example, vowel harmony typically does not apply to segments that are not contrastive for the harmonizing feature. While such a condition may simply be written into the rule itself (by excluding specific segments), this would, however, miss an important generalization: that the rule does not apply specifically in those cases in which it would create a segment that is not in the language's phoneme inventory. The absence of certain vowels in the inventory would therefore have to be stated twice: in the MSCs of a language, directly, and indirectly in the conditioning environment of the vowel harmony rule. This became known as the *duplication problem*, addressed by the principle of Structure Preservation, which states that the output of a rule must be a well-formed lexical representation (although this principle does not hold postlexically, where allophonic variation may be introduced). Other types of rule blocking include non-derived environment blocking, in which rules only apply to phonologically or morphologically derived sequences, or blocking in cyclic rule applications, for which the Strict Cycle Condition (Mascaró, 1976) was invoked, which protects material derived at an earlier cycle of rule application (and may thus block a rule from applying).

Developments in theories of phonological representations also introduce constraints on possible structures, most famously Autosegmental Phonology (Goldsmith, 1976a, 1990) and further developments thereof. Take, for example, the condition that association lines must not cross, or the Association Convention, which was designed to provide an automatic mechanism for associating elements on different tiers (although this became subject to parameterization, cf. Goldsmith, 1990). The idea that representations can function as constraints on possible and impossible structures and processes is probably taken furthest in theories of Feature Geometry (Clements, 1985; Clements and Hume, 1995; Halle, 1995; McCarthy, 1988; inter alia), which propose universal conditions on which features can function as sets in rules. The hope expressed in work on representations was that the rule component in phonology could be significantly

simplified, reduced perhaps to the automatic application of processes as a response to constraints on representations. This hope was aptly expressed by McCarthy (1988: 84): ‘if the representations are right, the rules will follow’.

Apart from constraints that were introduced as conditions on the architecture of a phonological theory, other constraints were introduced which made more direct reference to specific phonological structures. One such constraint is the Maximum Onset Principle (MOP) (e.g. Selkirk, 1982), which states that onsets take preference over codas in syllabification; this was assumed to be a universal principle, valid in all languages. Another famous constraint is the Obligatory Contour Principle (OCP) (Leben, 1973) which prohibits adjacent identical feature values and which received a central position in theorizing in the wake of Autosegmental Phonology. The OCP also points at a problematic status of such universal constraints; their universality may be disputed. Thus, while McCarthy (1986) promotes the status of the OCP to that of a universal constraint, Odden (1988) is much more cautious, showing that while the OCP is a powerful principle, it is not universal, citing counterevidence from some languages which appear to allow at least some OCP violations (although they may repair others).

This observation regarding the controversial universality of the OCP relates to another point. Given that there are recurring processes or SSCs in the world’s languages, which nevertheless are not universal (such as final devoicing, which is frequent but not found in many languages), what is their status in the theory? Are these constraints language-specific, emerging from rule interaction, or are they universal? And if they are language-specific, why are certain types of processes so frequent? We already saw the answer given by Natural Phonology: that there may be a universal inventory of possible rules – or constraints. These would perhaps be parameterized, switched on or off in individual languages. This line of reasoning is found, for example, in the Theory of Constraints and Repair Strategies (TCRS), to which we will turn in the next section, where we consider how the growing conviction that constraints occupy a central role in phonological theorizing led to the emergence of new phonological theories.

3. The Emergence of Constraint-Based Frameworks

It would go beyond the scope of this chapter to trace the history of phonological constraints further. It suffices to say that they had slowly but surely occupied a central position in phonological theorizing by the mid 1980s. This insight spurned a significant body of work looking into the question of how rule-based generative phonology could be reconceptualized more radically, to award constraints a more central position in the theory. This resulted in a diversification of the field and a number of new frameworks, none of which gained universal

acceptance, however, until the arrival of Optimality Theory (McCarthy and Prince, 1993b; Prince and Smolensky, 2004 [1993]) provided a framework that garnered not universal but widespread support again. Before discussing Optimality Theory in some detail, it seems pertinent to look at some of the alternatives that emerged slightly earlier, however, to illustrate the different directions in which phonological theorizing went, and to demonstrate the heterogeneity hidden beneath the label ‘constraint-based’.

What these theories share, despite all conceptual differences, is the assumption that constraints should be at the heart of phonological theory, making rules superfluous or demoting them to the status of repair strategies. We will now briefly look at three models which are alike in their focus on constraints but quite different in the actual formalism, the Theory of Constraints and Repair Strategies (TCRS) (Paradis, 1988), Government Phonology (GP) (Kaye et al., 1985) and Declarative Phonology (DP) (Bird et al., 1992). All three models enjoyed some popularity in the late 1980s and early 1990s, and Government Phonology remains to this day a fairly widely used framework, if somewhat outside the (US American) mainstream. I chose these three models as representative of three different approaches to a constraint-based paradigm. There are other constraint-based approaches that emerged around this time as well (e.g. Harmonic Phonology, Goldsmith, 1993; Grounded Phonology, Archangeli and Pulleyblank, 1994, both of which keep the notion of rules but heavily constrain their application through well-formedness conditions) but space does not allow to do all of them equal justice. The present selection is meant to be representative of different ways of theorizing at that time.

3.1 Constraints and Repair Strategies

TCRS (e.g. Paradis, 1988, 1996; Paradis and El Fenne, 1995) is perhaps best summed up as a close phonological ally of the Principles and Parameters approach to syntax. There are some universal principles (more on which presently), while constraints, formulated as markedness constraints which do or do not tolerate specific phonological structures, function as parameters which can be switched on or off. To return (once more) to final devoicing, the prohibition against final voiced obstruents would be one such constraint, with a language-specific setting (on in German or Russian; off in English or French). If a form violates an active constraint, this will activate a repair strategy. Repair strategies are single operations, much like the rules in Sommerstein (1974), which have been deprived of their context and are activated only to alleviate a constraint violation, although Paradis is adamant not to call them rules.

Final devoicing thus is motivated by a universal constraint against a [+voice] feature in final position, which can be active or inactive in individual languages

(as a parameter setting). The repair of this violation now is subject to a number of principles. One of them is the Minimality Principle which states that a repair takes place 'at the lowest phonological level to which the violated constraint it preserves refers' (Paradis and E1 Fenne, 1995: 188), which here is the feature. Another principle is the Threshold Principle (see, for example, Paradis, 1996): every repair must take place in maximally two steps. If it is impossible to repair the violation in two steps at the lowest level, the repair mechanism moves up one level (say, from the feature to the segment) to see if a possible repair can occur at that level. In the case of final devoicing, a repair is typically possible at the feature level. The reference to the phonological hierarchy thus also explains why consonant deletion or vowel insertion are typically not found as strategies to avoid a final voiced obstruent: the constraint refers to a lower level of the hierarchy (features). Deletion or insertion are thus only possible if a simple feature change would not solve the problem (e.g. if devoicing were to create an illicit segment). It is worth noting that to this day, Optimality Theory is dogged by the problem that a constraint violation could, in principle, invite a number of possible repairs, many of which, however, are robustly unattested. This has been dubbed the 'Too Many Repairs' problem (e.g. Lombardi, 2001b; Steriade, 2001b), and a uniformly accepted solution to this problem is still not available in Optimality Theory.

3.2 Government Phonology

Another model that is heavily influenced by generative syntactic theory is Government Phonology (GP) (e.g. Charette, 1991; Kaye et al., 1990; Kaye et al., 1985), which is modelled extensively after Government and Binding Theory, although it does look quite different from TCRS. First, GP proposes a new theory of representations, at the heart of which there is a basic distinction between onsets (O) and nuclei (N). In early work ('classical' GP), these can be branching (thus allowing for complex onsets or branching rhymes). Later work (see especially Scheer, 2004) dispenses with complex constituents, replacing O and N with C and V positions (known as 'strict CV'). A consequence of this is the existence of empty positions; any consonant cluster, for example, is separated by an empty V position. There are already empty positions in classical GP. Word-final codas, for example, are universally analysed as onsets followed by an empty nucleus. Below the segment, GP also abandons central ideas of classical generative phonology, in replacing distinctive features with so-called *elements*, which stand in hierarchical dependency relations, as the building blocks of segments (e.g. Harris and Lindsey, 1995; see also Botma et al., this volume).

Second, the well-formedness of these representations (e.g. motivating the occurrence of empty positions) is subject to constraints which are formulated as

principles, in analogy to principles in Government and Binding syntax. GP thus introduces the notions of government and licensing, used, for example, to define a phonological version of the Empty Category Principle, which can trigger phonological repairs. So, the licitness of word-internal empty nuclei (found after a 'coda' consonant) is subject to that empty position being 'properly governed' by another non-empty nucleus. Instead of positing explicit phonological rules, GP thus attempts to capture phonological generalizations by well-formedness conditions on representations. These are formulated as abstract principles of licensing, rather than being grounded in notions of markedness or phonetic naturalness, though. GP thus also abandons the idea of derivational depth. Since there are no rules, there is no rule ordering. Rather, 'processes apply whenever the conditions that trigger them are satisfied' (Kaye, 1995: 291), showing a similarity to the TCRS where processes also only apply in response to a constraint violation. For details, see, for example, Charette (1991), Kaye (1995), Kaye et al. (1990).

Final devoicing is motivated in GP by abstract licensing requirements rather than by a phonetically natural constraint prohibiting coda voicing which is stated directly in the grammar. The main gist of the analysis in Brockhaus (1991, 1995) is the following: 'codas' are onsets which are followed by an empty nucleus which licenses the onset. Empty nuclei themselves have to be licensed in order to be licit (and word-final empty nuclei are licensed qua principle), but segments which are licensed are no good licensors themselves. An empty nucleus thus is not as good a licensor of an onset as a properly filled nucleus, with the result that it can only license less complex segments. Voiced segments thus shed some of their content (the L-element which is responsible for voicing) in order to be licensed. Final devoicing thus follows from abstract requirements of licensing and crucially depends on the notion of empty syllable positions rather than from a substantive, phonetically motivated constraint which states directly that voicing distinctions in codas are disallowed.

GP has established itself as an influential framework in phonological theory. Much work within GP is dedicated to segment-zero alternations, which can be handled very successfully. Other areas of phonology (such as stress) still remain relatively under-researched. Nevertheless, GP enjoys a significant following to the present day, especially in Europe and Canada.

3.3 Declarative Phonology

A third constraint-based theory to be introduced briefly is Declarative Phonology (DP) (e.g. Bird et al., 1992; Coleman, 1995; Scobbie, 1997). Like GP, the theory is based on ideas in generative syntactic theory. However, it does not take its inspiration from Chomskyan models but from declarative, unification-based

models of syntax, especially Generalized Phrase Structure Grammar (GPSG) and Head-Driven Phrase Structure Grammar (HPSG) (GPSG, Gazdar et al., 1985; HPSG, Pollard and Sag, 1994). DP also does away with rules, but in addition it also is strictly non-derivational: there is no transformation of an underlying representation to a surface representation. Instead, strings are built monotonically, subject to constraints on string concatenation (as the unification of strings).

This statement probably needs a short explanation. There is no distinction between underlying and surface structure. Instead, the surface structure is built directly from the lexical representation, governed by constraints, typically formulated as if–then constraints (which could therefore be easily reconceptualized as rules). These constraints define conditions for unification by stating how two chunks of lexical representation can be concatenated. Notice, therefore, that like GP, DP employs a quite different notion of ‘constraint’. They are neither grounded in markedness, nor are they principles of phonological well-formedness, defined by abstract notions of licensing and government. Instead, they define positive conditions on how lexical material can be merged or unified by explicitly stating the set of possible representations, post-unification.³

This means that there are no phonological processes in the classical sense, as changes occurring between underlying and surface form because all that the constraints do is unify the atoms of lexical representations. This also means that structure-changing operations are impossible; only structure building is allowed, and whatever has been built, cannot be changed. While this captures non-structure-changing allophony straightforwardly (which would only add information on the phonetic realization of lexically contrastive units), neutralization is more problematic.

The example of final devoicing thus does not receive a straightforward analysis in DP as an outcome of unification because, in traditional terms, it involves a structural change from underlying [+voice] to surface [–voice]. The DP solution to this problem is to have two representations in a lexical entry of a word, disjunctively, one underspecified (such that it can receive the appropriate feature value in structure-building unification), one specified as a default option if unification does not provide a feature value (Bird et al., 1992; Coleman, 1995). Final devoicing – or indeed any kind of neutralization – then is not a feature-changing operation. Rather, the question of neutralization or contrast reduces to the question of what type of lexical representation (specified or underspecified) is selected by the set of constraints that build structure.

With the notable exception of Government Phonology, none of the proposals that emerged in the late 1980s and early 1990s gained wide popular appeal. The plethora of proposals abounding at that time shows, however, that there was a general sense within the phonological community that the time was ripe for a break with traditional generative phonology. The theory that would finally gain

mass support was Optimality Theory (McCarthy and Prince, 1993b; Prince and Smolensky, 2004 [1993]).

4. Optimality Theory

Optimality Theory (OT) marks a radical departure from other constraint-based theories of phonology, in a number of respects. First, the notion of rule or repair strategy disappears completely, as does a meaningful notion of phonological operation. All changes or operations are motivated solely by constraint interaction. Constraints are universal in this theory, but they can be violated, and rather than being parameterized (on or off), they are ranked with respect to each other, by significance. There are only constraints on surface structures (outputs); OT does not have an equivalent of MSCs, or constraints holding at intermediate levels of representation. In addition, there is no single representation which undergoes transformations from input to output. Instead, there are several output candidates which compete for selection as the optimal output, the form that violates the least high-ranked constraints, and this competition is conceptualized as the parallel evaluation of different candidate forms. And there is another, significant change to the conceptualization of constraints in OT: while in previous theories constraints were largely *markedness* constraints that impose (potentially universal) conditions on possible output configurations, they are complemented with a class of *faithfulness* constraints in OT, which militate against changes applying to the input. Differences between grammars then reduce to the relative ranking of this universal set of constraints, especially the relative ranking of markedness and faithfulness constraints with respect to each other.

4.1 A First Case Study

Let us illustrate the basic principles by returning to final devoicing, but with a typologically broader perspective in mind, from which there are essentially three types of languages: (i) languages that allow voiced and voiceless obstruents in all positions (like English); (ii), languages that allow only voiceless obstruents in syllable codas (like German; final devoicing languages); (iii) languages that do not allow voiced obstruents at all, that either have no laryngeal contrasts (like Hawaiian), or only contrasts involving aspiration or glottal constriction (like Korean, disregarding allophonic voicing).

Now assume three constraints: First, the familiar constraint against voiced obstruents in codas, abbreviated here as *CODA VOI. Second, a constraint against voiced obstruents in general, in line with the finding that voicing is marked in obstruents: *OBS VOI. Finally, assume a faithfulness constraint against changing

underlying voicing specifications: IDENT(voi). Now suppose a hypothetical input /bad/ and three potential outputs: [bad, bat, pat]. We can schematically draw up how each of these forms satisfies or violates the above constraints in a tableau, with violations marked by an asterisk '*':

Tableau 7.1 Voicing and markedness

/bad/	IDENT(voi)	*CODAVOI	*OBsvOI
[bad]		*	**
[bat]	*		*
[pat]	**		

All candidates violate at least one constraint. The faithful candidate [bad] violates the markedness constraints, the prohibition against voiced codas and the general constraint against voiced obstruents, which is in fact violated twice [b, d]. The bottom candidate [pat] satisfies these markedness constraints, but at a price: it violates the identity constraint twice by showing two discrepancies in the feature [voice] vis-à-vis the input. The solution to the problem that there are no inherently optimal candidates is to rank the constraints with respect to each other.

Consider the order in which the constraints are listed in Tableau 7.1 to be a ranking order. Then, IDENT(voi) is top ranked; hence, the first, faithful candidate [bad] is selected as optimal because it is the only candidate that satisfies the top-ranked constraint. Conversely, the third candidate [pat] with two voiceless stops is optimal if *OBsvOI is ranked highest because it is the only candidate satisfying that constraint.

A final devoicing grammar results from yet another reranking of these three constraints, namely by ranking *CODAVOI highest and leaving *OBsvOI at the bottom of the hierarchy. This will select the final devoicing candidate [bat] as the winner (see Tableau 7.2). Faithful [bad] is eliminated in the first round, because it violates the top-ranked constraint against voiced codas. Now only [bat] and [pat] remain in the competition. IDENT(voi) decides over who wins: as [bat] violates the constraint only once, while [pat] violates it twice, it is selected as optimal. (Note that reranking the bottom two constraints would yield an all-voiceless grammar again, selecting [pat] as the best candidate.)

Tableau 7.2 A final devoicing grammar

/bad/	*CODAVOI	IDENT(voi)	*OBsvOI
[bad]	*!		**
 [bat]		*	*
[pat]		**!	

Tableau 7.2 also illustrates the basic formalism of constructing OT tableaux. The leftmost column lists (i) the input and (ii) output candidates, marking the optimal candidate ✱. The remaining columns list the constraints, in descending rank order, and mark violations of these constraints by asterisks. All candidates are evaluated in parallel, by constraint, starting with the highest-ranked constraint. In each round, the candidates that violate a constraint are excluded from further consideration, until the set of candidates is reduced to one, the optimal candidate which is selected as the output form. If a constraint eliminates all remaining candidates, this constraint is skipped, and evaluation continues one level down. Solid lines between columns indicate that the constraints are ranked while dashed lines (as in Tableau 7.1) indicate that two or more constraints are not ranked with respect to each other.

The permutation of constraints in this mini-grammar yields a *factorial typology*, the typology of different grammars (yielding different outputs) that can be constructed from reranking the same set of constraints. In this example, three distinct grammars can be constructed (disregarding those that do not select a unique winner, like the initial set of unordered constraints), and these fit the typological observations: languages may not allow voiced obstruents anywhere, display final devoicing or tolerate voiced segments in all environments. The goal of yielding a close match between observed typological variation and computationally possible grammars, and hence languages thus occupies a central role in standard OT: the idea of a universal constraint set with a finite number of possible ranking permutations makes a clear prediction of what grammars can or cannot exist. A language with initial devoicing is predicted not to exist because such a candidate could under no ranking emerge as optimal, as Tableau 7.3 shows, now including the initial devoicing candidate [pad].

Tableau 7.3 The non-optimality of initial devoicing

/bad/	IDENT(voi)	*CODAVOI	*OBsvOI
[bad]		*	**
[bat]	*		*
[pat]	**		
[pad]	*	*	*

The initial devoicing candidate could never win because it violates all three constraints while the other three candidates satisfy at least one of them. This means that, no matter which constraint we rank first, [pad] will be eliminated in the first round of evaluation. In other words, [pad] is *harmonically bounded* by the other candidates – there is always a better option available. OT thus predicts initial devoicing not to occur in any natural language.

4.2 The Emergence of Unmarked Forms

The ranking of different constraints can yield a typology of possible grammars, thus languages. This in itself would be rather unremarkable, since other theories also try to achieve a match between the descriptive power of a grammar and attested typological variation. However, the notion of ranked constraints (rather than parameterized on/off switches) also explains why unmarked structures can emerge in languages even if they allow corresponding marked structures, because although there are optimal forms which violate some constraint, the grammar still preferentially picks candidates that do not violate that constraint; constraint violation always comes at a cost. A parameterized theory of constraints has no such mechanism.

Let us illustrate this point with an example from syllable typology. There are languages in which all syllables must be open, that is, where all syllables are of the maximally unmarked type CV. In OT, the optimality of CV syllables is expressed by the interaction of two constraints: *ONSET* (syllables have onsets) and *NoCODA* (syllables do not have codas), as shown in Tableau 7.4. If *NoCODA* is high ranked in a language, above segmental faithfulness constraints (here expressed as a blanket faithfulness constraint *FAITH*), this means that a potential input with a final consonant will be repaired, by deleting the consonant or adding a vowel, thus yielding an additional open syllable. (We ignore the question of whether this language favours deletion of a coda consonant over insertion of a final vowel.)

Tableau 7.4 Languages with only open syllables: *NoCODA* \gg *FAITH*

/CVC/	<i>NoCODA</i>	<i>FAITH</i>
[CVC]	*!	
 [CV]		*
 [CVCV]		*

Languages which allow codas, like English, in contrast, must have the opposite ranking. Faithfulness ranks low and therefore codas are allowed; both candidates [CV] and [CVCV] violate *FAITH*, which in English would be ranked above *NoCODA*. This, however, does not entail that English freely tolerates violations of *NoCODA*. We have already seen that pre-OT work on syllabification suggested a Maximum Onset Constraint, which states that syllabification of consonants into onsets takes precedence over syllabification into codas (see Section 2.2). Thus, a string /CVCV/ is universally syllabified as [CV.CV], never as [CVC.V]. This observation is a direct consequence of the presence of *NoCODA*. No matter how low ranked this constraint is, the candidate with the

fewest codas, while remaining segmentally faithful to the underlying form, will always be optimal, as Tableau 7.5 demonstrates. Although FAITH outranks NoCODA, thus prohibiting the deletion of consonants to generate codaless syllables, a codaless syllable is still the preferred option where faithfulness is not at stake. Therefore, the first candidate, in which the intervocalic C is parsed as an onset, is selected over the second candidate, where the consonant occupies the coda position of the first syllable; because of this, the second candidate incurs an additional violation of NoCODA.

Tableau 7.5 Intervocalic consonants are onsets

/CVCVC/	FAITH	NoCODA
CV.CVC		*
CVC.VC		**!
CV.CV	*!	

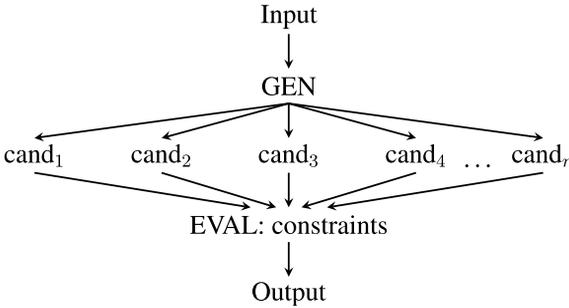
Traditional constraint-based theories, which assume that (an equivalent of) the NoCODA constraint is either switched on or off in a language, need an additional mechanism to explain why syllabification in languages that do allow codas nevertheless avoids creating them, for example, by reference to a separate onset maximization principle. The fact that no such additional mechanism is needed in OT shows one of the strengths of the idea of ranked constraints: even though marked structures may be allowed in a language, there is still the possibility for unmarked structures to emerge, embodied in the presence of (however low ranked) constraints. Such phenomena have been called The Emergence of the Unmarked (TETU) (McCarthy and Prince, 1994). Classically, they occur in contexts such as reduplication when the reduplicant shows a preference for unmarked syllable structure. However, we may stretch this idea somewhat and say that other phenomena where the unmarked structure will be selected as optimal, although marked structures are in principle permitted, are also a case of TETU.

4.3 Basic Principles of OT

This section will describe more closely the fundamental architectural properties of OT which have only been informally mentioned so far. Essentially, OT contains two central functions, the generator function GEN and the evaluator function EVAL. The job of GEN is to generate a set of candidate forms from an input (underlying form), while EVAL contains the ranked set of constraints that

will select the most harmonic candidate as the optimal output form, which is then fed to the phonetics interface. This is schematically shown in (9); each component will be discussed separately in turn.

(9) The architecture of OT: schematic view



The most important property of the input in OT is that there are no constraints on possible inputs, a property known as *Richness of the Base*. There are hence no MSCs or other comparable restrictions. A possible word in a language is determined solely by constraints on output forms. One could therefore posit hypothetical input forms for some nonce words, like /çøχ/ or /ktsqd/. EVAL would then select an output form that is a possible surface form of English, like [ʃɜ:k] or [kə'tɪsk]. The generalization that was mentioned in Section 2.1, viz. that all triconsonantal clusters in English are of the type [s]+[p,t,k]+liquid, is therefore also handled exclusively by output constraints (e.g. by a constraint stating that all triconsonantal clusters have to begin with [s]), not by conditions on inputs, such as underspecification.

The theoretical importance of this is that the analyst cannot exclude certain input forms by stipulation; any possible form has to be taken into consideration. The advantage of this principle is that it helps to defuse long-standing discussions in some languages about what the correct underlying representations of some forms are – in OT, the decision does not need to be made as long as all possible (or disputed) input forms map onto attested output forms. Otherwise, one may consider its practical value to be rather low – why would English have URs like /ktsqd/ in the first place? Still, the value of this principle is not just theoretical: in loanword adaptation, input forms can indeed be quite removed from the typical shape of morphemes in the borrowing language. In rule-based phonology, ad hoc rules have to be stipulated that repair a borrowing so that it conforms to the MSCs of the borrowing language and can become a lexical item. In OT, the same ranking that accounts for phonological patterns in the native vocabulary should be able to adapt and repair the phonological shape of new borrowings.

A principle quite similar in spirit to Richness of the Base governs the GENERATOR function, viz. *Freedom of Analysis*, which states that there are no constraints on candidate generation either. The candidate set is therefore potentially infinite. In practical terms this means that no possible form may be excluded from evaluation. Going back to the voicing typology in Section 4.1, none of the candidates under consideration may be excluded from the candidate set, including, theoretically speaking, the harmonically bounded initial devoicing candidate [pad]. However, in practice, only likely candidates are considered; candidates like [sli:p] or [fə'nɒlədʒi] for an input /bad/ are not seriously considered, since they involve multiple unmotivated violations of faithfulness constraints. Nevertheless, we assume that these forms are generated as well, but are simply eliminated very early on since they are harmonically bounded by other forms, and as such not interesting. For a computational solution to the theoretical problem of a potentially infinite candidate set, see also Riggle (2004), who proposes a Contenders algorithm that excludes harmonically bounded forms from generation and evaluation.

The notion of constraint evaluation has already been illustrated by the case studies in Sections 4.1 and 4.2. Let me summarize the main points here. First, the constraint set is assumed to be *universal*; all languages share the same set of constraints (although language-specific constraints may also exist, to account for idiosyncratic or morpheme-specific processes).⁴ Differences between languages are thus not (primarily) differences in the constraints they have but in the ranking of these constraints. Rankings have to be imposed because constraints are in conflict, the most basic type of conflict being between markedness and faithfulness constraints (about which more presently), and ranking resolves this conflict by imposing an order of relative importance on the constraints. These rankings moreover are *transitive*, which means that if constraint A outranks constraint B, for which we write the shorthand $A \gg B$, and B outranks C, then A also outranks C, precluding the possibility of circular rankings, where $A \gg B$ and $B \gg C$ but $C \gg A$. Related to this is the principle of *strict domination* which states that if one constraint outranks (or dominates) another constraint, nothing can reverse the ranking, including, for example, the number of constraint violations incurred by a candidate. So if $A \gg B$ and we have two candidates, one violating A once, one violating B five times, the second candidate will still win, or, to use proper OT terminology, will be more harmonic than the first candidate (although this view has recently been challenged by Pater, 2009, who proposes numerically weighted constraints, imported from Harmonic Grammar, which may cause such 'ganging-up' effects).

Finally, constraint evaluation is *parallel*, with several consequences. First, all candidates are evaluated in parallel, as discussed in Section 3.1. Second, there are no intermediate levels of representation. All changes to an input form apply

in one fell-swoop, unlike ordered rules, in which there is one change at a time (although serialist versions of OT have been proposed as well under the heading of harmonic serialism; see, for example, McCarthy, 2007, 2009). An input is therefore submitted to GEN only once, where any number of changes may occur, and EVAL only evaluates the product of the parallel application of all these changes. As we will briefly discuss in the concluding section, this has been one of the biggest criticisms of OT, since effects resulting from the ordered application of rules, such as opacity effects, cannot be modelled straightforwardly in OT. Before returning to this issue, we first discuss constraints in some more detail.

4.4 Constraints

We have already seen that there is a basic distinction between markedness and faithfulness constraints. Faithfulness, the true innovation of OT (markedness constraints having been well established in pre-OT work already), is commonly formalized as Correspondence (McCarthy and Prince, 1995) (early OT conceptualized faithfulness differently, as Containment; see McCarthy, 1993; McCarthy and Prince, 1993b; Prince and Smolensky, 2004 [1993], for examples). Correspondence constraints evaluate the degree and the quality of correspondence between elements in the input and a candidate (or in fact any two levels of representation). The classic definition of Correspondence is given in (10).

(10) Correspondence

Given two strings S_1 and S_2 , correspondence is a relation \mathfrak{R} from the elements in S_1 to those in S_2 . Elements $\alpha \in S_1$ and $\beta \in S_2$ are referred to as correspondents of one another when $\alpha \mathfrak{R} \beta$

(McCarthy and Prince, 1995: 262)

(10) makes it possible to define and evaluate different types of correspondence by constraints. The most frequently used constraints are MAX (no deletion), DEP (no insertion) and IDENT(F) (no feature value changes), defined in (11), (12) and (13), respectively.

(11) MAX Every segment of S_1 has a correspondent in S_2 .

(12) DEP Every segment of S_2 has a correspondent in S_1 .

(13) IDENT(F) Correspondent segments have identical values for feature F.

The terms MAX and DEP may seem confusing. They are motivated as follows: no deletion means that input material is MAXimally preserved in the output, while

no insertion means that all elements in the output DEPEND on the presence of corresponding elements in the input.

Note that correspondence need not be just correspondence between input and potential output forms (so-called IO-Correspondence). The same type of relation has been proposed to relate other levels of representation to each other as well, for example, correspondence between a base and a reduplicant (McCarthy and Prince, 1995) or Output-Output Correspondence (Benua, 1997). The latter type of correspondence can be used to account for cyclic and paradigm uniformity effects, where an optimal output form has to be phonologically similar to other, related forms. See also McCarthy and Prince (1995) and Kager (1999, chapters 5, 6) for an in-depth discussion of Correspondence.

Finally, note a special kind of faithfulness, so-called *positional* faithfulness (Beckman, 1998), which specifically protects salient positions in a word (e.g. word-initial positions, onsets). Positional faithfulness is used to explain the resilience of strong positions to undergo processes, or to determine the directionality of assimilation (e.g. formalizing the observation that codas tend to assimilate to onsets, rather than vice versa). The final devoicing example in Section 4.1 can be recast in a positional faithfulness gist, by having only one markedness constraint *OBSVOI and an additional faithfulness constraint IDENT(voi)/Ons, which protects voicing distinctions in onsets. Final devoicing would then not be due to the effect of a specific positional markedness constraint against voicing in codas but the effect of a faithfulness constraint which prohibits devoicing from occurring in onsets as well.

While faithfulness constraints are a formally strictly defined small class of constraints, the same cannot be said of markedness constraints. There are very many markedness constraints in the literature, and there is no comparable consensus over what the set of markedness constraints contains, although a core set of constraints seems uncontroversial. There are segmental markedness constraints which state which feature values or value combinations on segments are marked. We already encountered one such constraint in Section 4.1, *OBSVOI, which states that obstruents should be [-voice]. Others would be constraints against front rounded vowels, against back unrounded vowels, against pharyngeal consonants, against dental fricatives, against nasal vowels, and more generally, against any structure which can be considered marked in the sense of being typologically rare, acquired late or articulatorily difficult. These constraints can also be contextual, in which case they pertain not just to the relative markedness of a segment per se but also take into account the phonological context. While nasal vowels may be generally marked, for example, with few languages displaying contrast between oral and nasal vowels, vowel nasalization is unmarked before a nasal consonant. So while there may be a general constraint *V_{nasal}, there is also a more specific constraint *V_{oral}N that is violated by an oral vowel, but only if it directly precedes a nasal (see also Kager,

1999: 28 for an elaboration). Finally, markedness constraints do not just occur at the segmental level but at all levels of phonological representation. We already encountered two markedness constraints on syllable structure, *ONSET* and *NoCODA*. A third one would be **COMPLEX*, prohibiting complex (branching) constituents, like onset or coda clusters. Similarly, there are markedness constraints on stress assignment like **CLASH*, **LAPSE* etc.

While the general assumption is that all constraints are freely rankable, some sets of constraints come with a fixed ranking, corresponding to degrees of markedness on a scale; these are called ‘grounded constraint subhierarchies’ in Padgett (2001a). For example, consonantal place features are ordered on a universal scale, such that [dorsal] is more marked than [labial], which in turn is more marked than [coronal], yielding the hierarchy in (14) (Lombardi, 1997; see also de Lacy, 2006 for a different formalization of the same observation).

- (14) A place markedness hierarchy
**DORSAL* \gg **LABIAL* \gg **CORONAL*

Other hierarchies include vowel height (e.g. **HIGH* \gg **MID*; Beckman, 1998) and manner of articulation, with stops being less marked than fricatives, for example. A special type of universal constraint hierarchy derives from prominence scales (Prince and Smolensky, 2004 [1993]). One such scale is the sonority hierarchy. Prince and Smolensky show how a pairing of this scale with syllable constituents as constraints can elegantly account for cross-linguistic syllabification facts, including the long-standing puzzle of syllabification in the Imdlawn Tashlhiyt dialect of Berber (Dell and Elmedlaoui, 1985). In brief, the sonority scale translates into two constraint hierarchies, one for margins, which are less marked the lower the sonority is, and one for nuclei whose markedness decreases as sonority increases, thus attracting low-sonority segments into margins and high-sonority segments into nuclei. See Prince and Smolensky (2004 [1993]) for details.

A final type of constraint which has to be mentioned briefly is the class of alignment constraints, which transcend the markedness/faithfulness dichotomy. Put briefly, alignment constraints regulate the construction of prosodic structure and the concatenation of morphological structure. Their formal structure is the same: they refer to two types of constituents and the way their edges align. For example, the constraint *ALIGN*(Ft, L, Wd, L) states that the left edge of each foot should align to the left edge of a word; high ranked, this constraint will enforce left-to-right directionality in footing (see also Apoussidou, this volume). For a detailed discussion of alignment, see also Kager (1999) and McCarthy and Prince (1993a).

Further developments in OT have given rise to other classes of constraints which space does not permit to discuss in detail here. A central issue has been

the notion of constraint conjunction, including the question under which conditions constraints may be conjoined (see, for example, Lubowicz, 2005; Moreton and Smolensky, 2002), and new types of constraints, such as the ‘targeted’ constraints of (Wilson, 2001). The set of commonly assumed constraints is therefore still in a constant flux.

4.5 Capturing Phonological Processes

Phonological processes in OT are triggered by a ranking in which a markedness constraint \mathcal{M} outranks a corresponding faithfulness constraint \mathcal{F} and the input contains a structure penalized by \mathcal{M} . In this case, the faithful candidate, which contains a violation of \mathcal{M} , will be less harmonious than a candidate violating \mathcal{F} , under satisfaction of \mathcal{M} . We saw an example of this in Section 4.1: final devoicing is triggered by a high-ranked markedness constraint $*\text{CODA}\text{VOI}$; violation of lower-ranked $\text{IDENT}(\text{voice})$ is optimal vis-à-vis violating $*\text{CODA}\text{VOI}$.

The appeal of OT does not just come from its ability to capture a single phonological process, though. OT also accounts for different types of blocking behaviour that have been problematic in rule-based theories of phonology and that led to the introduction of ancillary principles like Structure Preservation and the Elsewhere Condition. To begin with Structure Preservation, a process may fail to apply if it were to create a segment which is not part of the inventory of contrastive segments in a language. We saw in Section 2.2 that having a blocking condition in the rule itself and an MSC against the offending segment creates a duplication problem since it states the same restriction twice, in different components of the phonology.

In OT, the absence of a segment $[x]$ from the inventory of a language is generally motivated by a high-ranked constraint against that segment: $*[x]$. Now add the above process-triggering ranking $\mathcal{M} \gg \mathcal{F}$, ranking $*[x]$ even higher (yielding a partial ranking of $*[x] \gg \mathcal{M} \gg \mathcal{F}$). Sometimes, satisfying \mathcal{M} may cause a violation of $*[x]$, in which case a violation of \mathcal{M} would be tolerated, and the process is blocked from applying. Let us consider a more concrete example, a language in which hiatus is resolved after high vowels by inserting a homorganic glide ([j] after [i], [w] after [u]), while hiatus is tolerated after non-high vowels because the language does not have non-high glides. This is essentially the situation in some dialects of English where *see[j] it*, *do[w] it* have epenthetic glides but *saw it* doesn’t (this does not apply to accents with ‘intrusive [r]’, of course; see Uffmann, 2007 for a discussion of the relationship between glide insertion and intrusive [r]). Hiatus resolution is triggered by the ONSET constraint. In the gist of the analysis of intrusive [r] in Krämer (2008), assume that the faithfulness constraint violated specifically by glide insertion is $*\text{MULTIPLECORRESPONDENCE}$, which prohibits an input segment to have several output correspondents (the

glide being a second faithful correspondent of the input vowel). Now, given the ranking $\text{ONSET} \gg *MC$, we can say that glide insertion will be tolerated if the payback is satisfaction of the onset requirement. However, a corresponding non-high glide would violate a top-ranked constraint against non-high glides, say $*G_{[-hi]}$. We know this constraint is top ranked because the language does not allow non-high glides independently. Now the same constraint can also actively block the creation of a non-high glide. A violation of ONSET is thus optimal vis-à-vis the creation of an illicit glide.

In OT, there is only a small step from such blocking behaviour to effects of disjunctive ordering and the Elsewhere Condition, by introducing a second faithfulness constraint, call it \mathcal{F}' , which is ranked between \mathcal{M} and \mathcal{F} , giving a ranking of $*[x] \gg \mathcal{M} \gg \mathcal{F}' \gg \mathcal{F}$. Now if satisfaction of \mathcal{M} through violation of \mathcal{F} (the standard repair) would entail the violation of higher-ranked $*[x]$, violation of the other faithfulness constraint \mathcal{F}' may not. Thus, a different process is triggered if the standard process were to create an illegal structure.

To make this point more concrete, let us return to the hiatus resolution mini-grammar and assume a language in which hiatus is also resolved after non-high vowels, but by means of a different process, say, glottal stop insertion. (This is essentially the situation in German where a corresponding glide is found after the high vowels in *Pil[j]ano*, *Hy[ɥ]äne* 'piano, hyena', but a glottal stop is inserted elsewhere: *O[ʔ]ase*, *cha[ʔ]otisch* 'oasis, chaotic'.) Assume that the inserted glottal stop violates DEP(Lar) , a constraint against the insertion of laryngeal features. This constraint is ranked above $*MC$ but below ONSET ; glide insertion is thus optimal vis-à-vis glottal stop insertion unless high ranked $*G_{[-hi]}$ were to be violated. Then, ONSET can then still be satisfied, by violating DEP(Lar) .

An in-depth discussion of different types of blocking behaviour ('do unless', 'do only if') is found in Prince and Smolensky (2004 [1993]). For the moment, this short and rather abstract overview should suffice. The important point here is that blocking, which was a notorious problem for rule-based accounts, is formalized quite elegantly as constraint interaction in OT, which may be one of the main reasons for the immediate success this model enjoyed upon its inception.

5. Discussion

This chapter has provided a short introduction to Optimality Theory, based on a historical perspective. We chose this historical perspective because there seems to be a tendency to juxtapose OT with rule-based phonology, when, in fact, the distinction is much blurrier than that juxtaposition would suggest. Instead, the development of OT is an endpoint of a development that had been taking place for a while, a development in which the power of rules was

increasingly reduced while constraints, interpreted here in a more inclusive manner, assumed a more and more central position in phonological theorizing. What sets OT apart from other theories developed around that time is therefore less its focus on constraints itself but rather the radical way in which it disposed of some of the heritage of classical generative theory. Rules were not replaced with (formally still related) repair strategies (as in the TCRS) or if-then constraints (as in DP); instead, the idea that a process as such has a formal representation was abandoned. In OT, processes became an epiphenomenon of constraint interaction.

The relatedness of OT to other constraint-based theories also means that some of the criticism levelled against OT has to be seen within a wider context. As space does not allow for a comprehensive discussion of the perceived shortcomings of OT, I want to point out just one problem that has repeatedly come up in the OT literature and that many phonologists perhaps see as the Achilles heel of that theory, namely opacity effects. Cases of over- and underapplication are a problem for every theory of phonology that relies on surface constraints and/or the rejection of ordered rules. Opacity has been used as the biggest argument in favour of rule-based phonology since it suggests that processes do in fact apply in some order, and that different types of interaction between processes can be conceived of as the reordering of rules (see Bye, this volume, for a detailed discussion of rule ordering). It should be clear from the foregoing, however, that OT is not alone in its problems with dealing with opacity. Other surface-oriented theories like Natural Phonology suffered from the same problem, and all theories that emerged in the 1980s, with their rejection of extrinsic rule ordering, exploiting the idea that processes instead apply whenever their conditions are met, do so as well. TCRS, Declarative Phonology and Government Phonology thus share this potential shortcoming with OT. Reactions have been varied. Kaye (1995) simply rejects the idea that rule ordering phenomena exist; under- and overapplication of processes can occur in Government Phonology but are predicted to be the natural consequence of process interaction in this model. In other words, in environments where opaque interactions are attested, transparent process interaction is predicted not to exist. The details of this, however, do not always seem to be clear.

In Optimality Theory, under- and overapplication are generally problematic, given the strictly surface-based orientation of this model (though see Bakovic, in prep, for a qualification and for the demonstration that the common equation of counterfeeding and counterbleeding rule orders with the underapplication and overapplication of processes is far too simplistic). Several responses have come from within OT to address this problem. Some reject the existence of opaque interactions, reducing them to allomorph selection (e.g. Green, 2006; Sanders, 2003) while others have proposed additional machinery to deal with opacity phenomena, such as Sympathy Theory (McCarthy, 1998), Stratal OT

(Bermúdez-Otero, in prep.), Turbidity Theory (Goldrick, 1998), and others. A recent development has been to bring serialism back into OT, either as the serial evaluation of a derivation in which only one change occurs at a time until the system can no longer be optimized (harmonic serialism), or as Candidate Chain theory (McCarthy, 2007, 2008a, 2009). This also shows that OT is still constantly changing, in trying to maintain the advantages of a surface-oriented constraint-based theory while at the same time incorporating some of its criticisms and addressing its empirical shortcomings.

Constraint-based theories of phonology, including but not limited to OT, may therefore offer many conceptual advantages over rule-based theories of phonology. A number of these advantages have been addressed here, for example, the insight that many linguistically relevant generalizations can only be made from the perspective of surface strings, and the ability to provide a unified explanation to what were previously thought to be disparate processes. They may also suffer from shortcomings owing to their rejection of derivation depth. And so the debate continues to the present day although the pendulum has, at least since the 1980s, been swinging clearly in the direction of constraint-based models.

6. Notes

1. The generalization for German is actually more complex since all syllable codas are affected; for illustrative purposes, this simplified view should suffice, however. Note also that some research suggests that this neutralization to voiceless may be phonetically incomplete; see van Oostendorp (2008) for examples and a discussion regarding the relevance of this for phonological theory.
2. We will not be concerned with glides here since they are [-consonantal].
3. An anonymous reviewer suggests that DP, therefore, does not make a formal distinction between representations and constraints.
4. The standard assumption is that the set of constraints is provided by Universal Grammar (e.g. Prince and Smolensky, 2004 [1993]); however, constraints may also be emergent and learnable, on the basis of phonetic knowledge, but would still be universal because they are functionally grounded in articulation and perception (e.g. Ellison, 2000; Hayes 1996).

8 Phonetics-Phonology Interface

Silke Hamann

Chapter Overview

1. Introduction	202
2. Recurring Topics of the Phonetics-Phonology Interface	204
3. An [a] is an /a/ is an [a] is an \uparrow a \downarrow ? The Question of Representations	208
4. An Explicit Model of the Phonetics-Phonology Interface	216
5. Conclusion: No Theory-Neutral Interface	222
6. Notes	223

1. Introduction

To define the interface of phonetics and phonology, we first have to establish the topics of the two linguistic disciplines that are involved. While phonology deals with abstract, discrete categories encoding language-specific differences in meaning, phonetics is concerned with gradient, continuous, physical representations encoding articulatory and auditory details (see, for example, Pierrehumbert, 1990 and Cohn, 1993 for detailed discussions). The main question for the phonetics-phonology interface is how these two types of representations can be connected: how do we map discrete, abstract categories onto a phonetic parametric space and vice versa? This question is often restricted to one direction of processing, namely the production side, where it is known under the name of ‘phonetic implementation’: how are phonological structures that only exist in our brain converted into something in the real world?

Until the mid 1970s, the view prevailed that every phonological category (i.e. every segment, feature, syllable etc.) has an invariant characteristic in the acoustic signal or in the articulation. This invariant trait was considered to be language independent. The mapping from abstract phonological material onto the discrete phonetic realization was thus believed to be universal and

automatic. The search for these invariant, language-independent phonetic characteristics of phonological categories caused a massive increase in experimental methods, which eventually gave rise to the field of *Laboratory Phonology* (see Cho, this volume). However, finding invariant traits turned out to be more problematic than anticipated.

Early cross-linguistic studies on the acoustics of segments showed already that there are no universal correlates of phonological categories, see, for example, the work on voicing distinctions by Lisker and Abramson (1964, 1970, and later). Such findings had a huge impact on the conception of the phonetics-phonology interface (see the discussion in Fromkin, 1975 and Keating, 1988). Instead of an automatic process, phonetic implementation turned out to be something that has to be learned and that makes up part of our grammatical knowledge. Later studies revealed that even within a language there is no single invariant characteristic of phonological categories. In addition to inter- and intra-speaker variation, experimental work showed that the realizations of a category depend largely on context and prosodic position of the category, on the lexical frequency of the word in which it occurs, and on gender, regional background and social affiliation of the speaker (for literature on all these factors see the volumes on *Laboratory Phonology* starting with Kingston and Beckman, 1990).

How can speakers/listeners deal with all this variation? And is this large amount of variation reconcilable with the notion of an abstract representation at all? Instead of an answer to the mapping between phonology and phonetics, the experimental studies provided us with more questions. This situation led Kingston (2006: 401) quite recently to claim that there is still no consensus on what constitutes the phonetics-phonology interface. A scholar new to the field might wonder how this is possible after such a long time of extensive studies. The answer proposed in this article is that the nature of the interface cannot be unearthed by experimental studies alone. It depends to a considerable part on the theoretical assumptions we make, and on the aim we have in mind with our phonological and phonetic description. This has been observed already by Ladefoged (1972) almost 40 years ago, who summarized that there is not one but *many different* phonetics-phonology interfaces, depending on the aim of the studies and the assumptions following from it.

The present chapter is structured as follows: Section 2 shortly introduces three topics that can be considered essential for the interface, namely phonological features, the distinction between phonetic and phonological processes, and phonetic grounding. Section 3 deals with assumptions on phonological and phonetic representations, their possible connection, and how these assumptions shape the interface. A model with an explicit formalization of the interface and its predictions are introduced in Section 4. Section 5 concludes.

2. Recurring Topics of the Phonetics-Phonology Interface

In this section, we look at three topics that are important for the phonetics-phonology interface and can be found recurrently in the literature on the interface. These are as follows: phonological features (Section 2.1), possible distinctions between phonetic and phonological processes (Section 2.2), and the phonetic grounding of phonological processes (Section 2.3). They are relevant to the definition of the interface for the following reasons: while features are often considered to constitute the phonetics-phonology interface per se, the other two topics have a considerable impact on the distinction between phonetics and phonology and therefore on the connection between the two modules.

2.1 Phonological Features

Phonological features are the smallest unit in phonology (next to moras with their comparatively restricted use). Features are defined as having direct correlates in articulation, acoustics or audition (see Botma et al., this volume). Very often, they seem to be nothing more than a phonetic characteristic. The feature [labial], for instance, is shared by segments articulated with the lips. For this reason, features can be considered the interface between phonetics and phonology: They connect a phonological category (the segment) with a phonetic realization (its articulation, its percept or its acoustic form). The feature set that is most commonly used in present-day phonology goes back to the one proposed by Chomsky and Halle (1968) and has almost only articulatory correlates. This set has been modified by Clements (1985) to mirror the dependencies between articulators (e.g. coronals can be articulated either with the tongue tip or blade), resulting in a more realistic model of concrete articulations associated with phonological categories. Alternatives to these articulatory features and thus a purely articulatory definition of the interface are proposed, for example, by Jakobson et al. (1952), who introduced a feature set with acoustic, auditory and articulatory correlates, or the elements proposed by Harris and Lindsey (1995), which are unary features with acoustic definitions.¹

Besides their function as phonetic interpretation of segments, features are used to group segments in the description of phonological processes: segmental classes that undergo a process or form the context of a process are referred to by features. The class of sounds undergoing final devoicing in Dutch and German, for instance, are obstruents, defined by the feature [-sonorant].

Three problems arise for the phonetics-phonology interface with respect to the phonetic interpretability of features. First, experimental studies show

that phonological categories, including features, do not have one invariant phonetic trait (see the discussion in Section 1). The direct phonetic interpretation of features is thus restricted and has to be supplemented by additional information on their context- and language-specific realization. If features constitute the interface, a listing of all additional language-specific correlates of features seems necessary for a complete description of the interface. Second, features can be defined in articulatory, acoustic and perceptual terms, and while articulatory specifications seem more straightforward in the production (the traditional 'phonetic implementation'), a perceptual or acoustic specification seems to be more adequate for the perception process. If features constitute the interface, the question has to be answered whether the interface is restricted to production, perception or includes both (this point is taken up again in Section 3.4). Third, the phonetic traits of features cannot always be united with the function of phonological grouping. Sometimes a group of segments in a language forms a class in a phonological process but does not share a phonetic characteristic (see, for example, Mielke, 2004 for examples of such 'unnatural' classes). If features are assumed to come with their phonetic characteristics and thus provide the phonetics-phonology interface, then a different set of features not based in the phonetics seems necessary to account for unnatural phonological classes.²

In sum, referring to phonological features is not sufficient to fulfil the task of defining the phonetics-phonology interface. If features are assumed to constitute (at least part of) the interface, and it seems reasonable to connect phonetic realizations to the smallest phonological units, our assumptions on these phonetic realizations and on the use of features have to be made explicit. The questions raised here indicate how this could be done.

2.2 How Can We Distinguish between a Phonetic and a Phonological Process?

The decision whether a certain process is phonological or phonetic is not always as easy as it seems. Word-final obstruent devoicing, for instance, has been traditionally considered a phonological process: Obstruents (the class of [-sonorant] segments) change their specification from [+voice] to [-voice] in word-final position, resulting in neutralization between voiced and voiceless obstruents in this position. More recent findings show that this process of devoicing is not complete. There are small acoustic differences between devoiced and underlyingly voiceless obstruents which can be detected by native listeners, see, for example, Port and O'Dell (1985) for German and Warner et al. (2004) for Dutch. But if the output of this process is small, traceable remnants of voicing, can it still be considered phonological? The answer to this question depends on our

definition of phonology. If we consider phonology to operate on abstract and discrete representations (such as the feature [±voice]), as we have defined it at the beginning of this chapter, then an incomplete neutralization cannot be part of phonology but must be assigned to phonetics. The slightly disturbing question that arises is how many of the processes that have been traditionally considered phonological (as neutralizations) and have been described in textbooks of phonology really are phonological, or are only lacking the phonetic evidence to show their gradual nature.

In an alternative conception of phonology that allows gradient phonological representations, processes with non-discrete outputs can be considered phonological. Flemming (2001), a proponent of such a non-discrete approach to phonology, mentions a further possible problem for the strict division between phonological and phonetic processes in the traditional view: the two types of processes often seem to be the same, and cause a reduplication of description in phonology and phonetics. Flemming mentions phonetic coarticulation and phonological assimilation processes as an example for such a reduplication. Both involve the change of one segment caused by an adjacent segment; in one case this change is gradual, in the other case discrete. Flemming argues that treating both as caused by the same type of restriction (an effort-avoidance constraint) in the phonological grammar simplifies the grammatical description.³ This non-discrete approach to phonology and possible problems with it are discussed further in Section 3.2.

We can observe that only explicit criteria of what the output of a phonological process is allowed to look like helps us decide whether a process is phonological or phonetic. The way we define our phonological module thus determines what we consider to be part of phonetics. This in turn determines our conception of the phonetics-phonology interface, as we will see in Section 3.

2.3 Phonetic Grounding of Phonological Processes

A third topic that is mentioned recurrently in studies on the phonetics-phonology interface is phonetic grounding. The urge to seek phonetic motivations for phonological processes is far from new (see, for example, Passy, 1891) but has not lost its fascination, see more recent studies on phonetic explanations, for example, in *Natural Phonology* (Stampe, 1972a; Donegan and Stampe, 1979), in *Grounded Phonology* (Archangeli and Pulleyblank, 1994), and in the work on sound change by Ohala (e.g. 1981, 1983, 1992 and following). The reason why researchers try to phonetically explain phonological observations is the fact that large numbers of unrelated languages often show similar phonological behaviour due to phonetics. We find for instance that many languages with retroflex segments avoid sequences of retroflexes and high front vocoids, which

has been explained as the avoidance of sequences with extremely dissimilar tongue positions (Flemming, 2003; Hamann, 2003). The same articulatory motivation has been applied to the cross-linguistic avoidance of sequences with a rhotic and the glide /j/ (Hall and Hamann, 2010). But there are also perceptual motivations, which are more often employed to account for the preference of whole segment inventories. The best-known example is the so-called dispersion effect in vowel inventories. Five-vowel inventories, for instance, mostly consist of [i e a o u], which has been explained by the fact that these five vowels are perceptually most distinct, that is, are least often confused with each other (Ladefoged, 1975: 235–238; Lindblom, 1986).

Phonetic explanations like these have often been located at the phonetics-phonology interface. Scholars, however, differ on the question whether such explanations should be treated in linguistic theory at all. Let us look first at the approach taken in *Optimality Theory* (Prince and Smolensky, 2004 [1993]; McCarthy and Prince, 1993b, henceforth OT). Most OT approaches use phonological markedness constraints that incorporate the phonetic motivation directly into the phonological account. The avoidance of sequences with a rhotic and the glide /j/, for instance, can be captured in a constraint *rj, which is based on the articulatory complexity of this sequence. Such markedness constraints have the advantage that the phonetic explanation is transparent in the phonological description, and that the typological variation we find in the languages of the world can be accounted for with a difference in ranking between the relevant markedness and faithfulness constraints (what is known in the OT-literature as ‘factorial typology’). However, such accounts duplicate phonetic information in phonology (see the criticism voiced by, for example, Blevins, 2004; Hume, 2004/to appear; Ohala, 2005; Haspelmath, 2006). Others therefore propose that the phonetic basis of phonological processes should not be part of the linguistic theory at all, but should be sought ‘in a technically “extralinguistic” domain, that of historical linguistics’ (Hale, 2000: 241–242). Haspelmath (1999) and Blevins (2004) also argue for an account of the emergence of typologically similar patterns via diachronic adaptation, as given in Ohala’s work (1981 and following). What these latter studies lack is a formal account of how exactly such diachronic changes occur, that is, how the phonetics influences the acquisition process of a child in such a way that she acquires a phonological grammar that differs from that of her parents.

What the two approaches seem to have in common is the assumption that phonetic grounding can only be formalized in the phonological module of the grammar. Their assumptions on the nature of phonology guide their solution: if phonology is purely about discrete units, then we cannot formalize the phonetic grounding of processes in phonology (e.g. Ohala, Blevins, Haspelmath). If our phonological theory allows phonetic principles in the shape of markedness

constraints, on the other hand, then we can formalize phonetic grounding in phonology (traditional OT accounts). But why does phonetic grounding have to happen in the phonology? If we have an explicit definition of the phonetics-phonology interface, then the connection between the phonological module and the phonetic one can account for the phonetic motivation of phonological processes without duplicating phonetics in the phonology or leaving out the phonetic motivation at all (a proposal along these lines is described in Section 4).

Again an exact definition of what belongs to the phonological module and what to phonetics seems essential to the approaches presented here. In the following section we look at the phonetics-phonology divide in more detail, namely, at the representations that can be postulated in both modules and their possible connections. These connections, and the formalization of them, constitute the phonetics-phonology interface.

3. An |a| is an /a/ is an [a] is an † a † ? The Question of Representations

The focus of this section is the question of how the number and types of representations that we assume in our grammar model influence the definition of the phonetics-phonology interface. We can discern roughly three views on phonetic and phonological representations. According to the first view, phonetic representations are very detailed and there are no permanent phonological representations, see Section 3.1. The second view assumes that the two types of representations are fairly similar, as discussed in Section 3.2, and the third that they are incommensurable, see Section 3.3. Section 3.4 deals with two questions that arise if we assume that phonetic representations are different from phonological forms: how many phonetic forms can we discern, and how do they connect to the phonological representations.

3.1 No Stored Phonological Representations

The most extreme view on phonological representations is found in *Exemplar Theory* (Johnson, 1997, 2005; Pierrehumbert, 2001; the following is a crude summary, for a complete overview the reader is referred to Silverman, this volume). Exemplar Theory originated in cognitive psychology as an approach to concept learning. It has been applied to phonology due to the findings in Laboratory Phonology that the acoustic signal varies to such a large amount that it seems unreasonable to assume speakers have stored abstract representations and the means (rules or constraints) to derive them. In addition, it allows the connection

between phonetic realizations and gender, social status and other non-linguistic information on the speaker. Exemplar theorists do not postulate a separate phonological component of the grammar but assume instead that phonological generalizations are formed on-line across the stored phonetic exemplars when necessary. This process of generalization is described in the following quote from Beckman (2003):

It is abstracted away from sensory experience only to the extent that two utterances with values that are close enough to be indistinguishable along some dimension of the phonetic space will be registered as two instances of the same thing along that dimension. (p. 79)

This minimal view of representations is difficult to reconcile with the notion of a phonetics-phonology interface, because if there are no phonological representations we cannot formalize a connection between them and the phonetic forms. Exemplar Theory is therefore not further elaborated in this chapter.

3.2 Phonological and Phonetic Representations Are Commensurable

What does it mean to assume that phonetic and phonological representations are fairly similar? Hyman (2001) summarizes this quite fittingly by writing that both ‘phonetics and phonology utilize a common alphabet’ (p. 143) because, as he goes on to argue, they use the same descriptive terms such as ‘voiceless stops’ and ‘nasalized vowels’. The view that phonology and phonetics use fairly similar representations can be found in a large number of phonological studies.

Chomsky and Halle (1968; henceforth SPE), for instance, propose that phonological rules change binary phonological features stepwise into non-distinct, scalar specifications (what they term phonetic features). They assume that intermediate stages have ‘representations of a highly mixed sort’ (p. 66). These assumptions imply that phonological and phonetic representations are in some way comparable. According to SPE, the phonetics-phonology interface is then the point where abstract representations are complemented with (or replaced by) gradient details. This view is quite prominent in later generative work (e.g. in the assumption that features provide the phonetic realization of segments language-independently, see Kenstowicz, 1994 for an overview).

The idea that phonetic and phonological representations are commensurable can also be found in approaches that make no distinction between phonetic and phonological surface representations, and incorporate phonetic details directly into phonology. Such models are summarized here as *minimal functional approaches*,⁴ and the representations and their relation(s) look as given in Figure 8.1.

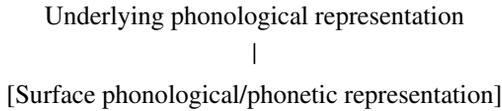


Figure 8.1 The two representations in a minimal-functionalist approach

Three arguments have been brought forward in support of this approach. First, there is allegedly no clear-cut distinction between phonetic and phonological representations. This point is illustrated with a quote from Flemming (2001: 8f.):

Phonetics and phonology are not obviously distinguished by the nature of the representations involved, [. . .] most of the primitives of phonological representation remain phonetically based, in the sense that features and timing units are provided with broadly phonetic definitions. This has the peculiar consequence that sound is represented twice in grammar, once at a coarse level of detail in the phonology and then again at a finer grain in the phonetics.

This quotation exemplifies the idea familiar from SPE that phonetic and phonological representations are commensurable. But instead of ‘duplicating’ information on speech sounds in phonetics and phonology and focussing on the phonological part as done in SPE and later generative work, Flemming proposes that we should treat phonetic and phonological representations as the same, implying that both are subject to the same grammar component. Their only difference then lies in their role: phonological representations are those distinctive contrasts that are chosen language-specifically amongst all possible phonetic contrasts.

The second argument brought forward for an incorporation of phonetic details into phonological representations is that phonetic and phonological processes often seem to be the same, see Section 2.2. Third, the phonetic grounding of phonological processes should be visible in the phonology, recall Section 2.3. In OT accounts, markedness constraints can refer to phonological structure (e.g. *CODA) or to phonetic detail (e.g. the effort minimization constraint *RETROFLEX in Flemming, 1995/2002, or the MINDIST constraint by Padgett, 2001b requiring perceptual dispersion). Such approaches can be found in the work by, for example, Flemming (1995/2002, 1997, 2001, 2003, 2004), Padgett (2001, 2003b, 2003a), Kirchner (1997, 1998/2001, 1999/2004) and Steriade (1995a, 1997, 1999c, 2001a, 2001b) (see Gordon, 2007 for an overview and Hale, 2000 for criticism).

In such a minimal-functionalist approach it is impossible to define the phonetics-phonology interface, since there is no clear-cut distinction between phonetic and phonological representations, and both are handled in one grammar component. In the following section, we see approaches that distinguish

phonetic from phonological representations, where the phonetics-phonology interface is thus the mapping between these representations.

3.3 Phonetic and Phonological Representations Are Inherently Different

As illustrated in Section 3.1 with Hyman's quote, the idea that phonetic and phonological representations are commensurable is largely based on the fact that both forms are represented with the same alphabet, namely the IPA (International Phonetic Alphabet) notation. In phonology, more precisely when depicting phonological forms, the IPA symbols stand for abstract representations, that is, segments, which themselves can be treated as shortcuts for feature bundles. In phonetic forms, IPA symbols can stand for articulatory gestures (e.g. a raised tongue tip), acoustic information (e.g. a high second and third formant), or for perceptual cues (e.g. energy in the mid region of the spectrum). Phonological accounts, however, often fail to give an explicit listing of what notations like /a/ and [a] stand for.⁵ As a result, researchers often seem to forget that these forms cannot be the same: an abstract phonological form with features and syllable boundaries cannot at the same time be a concrete articulatory gesture or an acoustic form. We will look now at alternative approaches with disparate representations for phonetics and phonology, as has been argued for by, for example, Pierrehumbert (1990, 377ff.).

If a strict division between abstract phonological representations and detailed phonetic representations is made, then it follows that phonological processes (in the phonological module) can only be described as a mapping between phonological representations, that is, of phonological onto phonological ones. The input to this mapping is the lexical form with segmental and featural representations. The output, the surface phonological representation, involves predictable phonological material such as syllable boundaries or feet. It further follows that, if one wants to include phonetic representations (such as articulatory gestures or other phonetic details), they have to be assumed *in addition* to the phonological ones. These minimally three representations are illustrated in Figure 8.2.

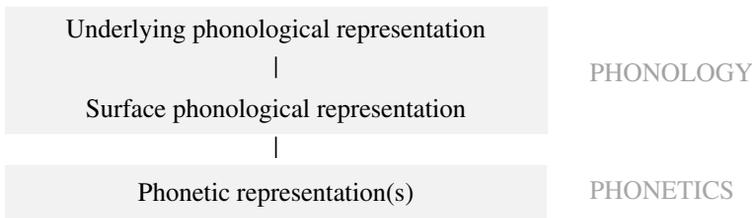


Figure 8.2 Phonological and phonetic representations in a modular approach

In the literature, we can find several phonological approaches that restrict themselves to the description of the two phonological representations. Since the present article is on the phonetics-phonology interface, we are, however, interested in those phonological approaches that include phonetic representations and make explicit statements on the connection between these different types of representations. There are two approaches in the phonological literature that fulfil these criteria. In the first, the connection between the surface phonological representation and the phonetic representation(s) is assumed to be automatic, in the second it is language-specific.

The first view is adopted by Hale (2000) and Hale and Kisko (2007). In order to distinguish three representations, they do the following. The traditional notation with slashes, as in /a/, are employed for an underlying representation. Brackets, as in [a], are used for what they call 'phonetic representation', which is the output of the phonological grammar or the form 'at the end of a derivation' (Hale, 2000: 243). Additionally, they assume a 'phonetic realization', which they define as the acoustic impression or the articulatory realization of a segment, and transcribe with little human bodies, for example, ♁ a ♁ . The latter is also called 'bodily output'. The mapping between the phonetic representation and the bodily output is performed by what they call 'transducers'. Transducers convert one type of representation into another, namely a featural, phonological one into a phonetic one and vice versa. In this respect, transduction differs from phonological computation, where only the same representations are involved. Hale and Kisko (2007) differentiate two types of transducers, an auditory and an articulatory one. The auditory transducer converts the percept into phonological features whereas the articulatory transducer converts phonological features into what they term 'a gestural score'. The exact working of these transducers, that is, a formalization, is not provided. The two transducers are innate and invariant, and 'identical in all humans' (Hale and Kisko, 2007: 84). The connections between the surface phonological representation and the phonetic representations are thus automatic. Together with Hale and Kisko's assumption that phonological features are drawn from a universal set, this results in the same phonetics-phonology interface for every language; the only task for researchers working within this framework on the interface is then to describe the automatic mapping between surface form (or more specific: phonological features) and the phonetic form(s).

The second view, where the connection between surface phonological structure and the phonetic form is learned and therefore language-specific, is proposed by Boersma in his *Functional Phonology* model (1998), and in the follow-up model of *Bidirectional Phonology and Phonetics* (Boersma, 2007, 2009). As we saw already for Hale and Kisko above, the use of a third representation requires the introduction of an additional notational convention. Boersma solves this by using pipes for the underlying phonological form, as in |a|,

slashes for the surface phonological representation, as in /a/, and brackets for the phonetic form(s), as in [a]. As in Hale and Kissock's approach, both directions of processing are described in this model, and for this purpose the phonetic form is further distinguished into an articulatory form and an auditory form. The latter is directly connected to the phonological surface form, and this connection constitutes the phonetics-phonology interface. It is arbitrary, which means that the learning child has to acquire the relevant perceptual cues for each segmental category. Boersma employs the same formalism for the connection between auditory and surface form as he uses for the connection between surface and underlying form (ranked OT-constraints, see Section 4 for details). This implies that both phonology and phonetics are treated as part of the grammatical knowledge.

3.4 How Phonological and Phonetic Representations Can Connect: Possible Views of the Interface

Both proposals discussed in the previous section argue for phonetics and phonology as two separate modules, and the mapping of a phonetic representation onto a phonological one as the locus of the interface. Both Hale and Kissock's and Boersma's proposals also include two directions of processing: a comprehension direction (the task of the listener) and a production direction (the task of the speaker). This contrasts with the majority of phonological studies, who focus on the production process (see the overview in Boersma and Hamann, 2009a). Even most phonologists working on the perceptual basis of phonological processes, such as Hume and Johnson (2001), Flemming (2004, 2005), Steriade (1995b, 2001a) and Wright (2001), assume or explicitly claim that speech perception informs phonology but lies outside the scope of phonological theory (exceptions are Broselow, 2004a and Pater, 2004). This view of phonology restricts the phonetics-phonology interface to the connection between surface and articulatory representation, as illustrated in Figure 8.3. Note that the

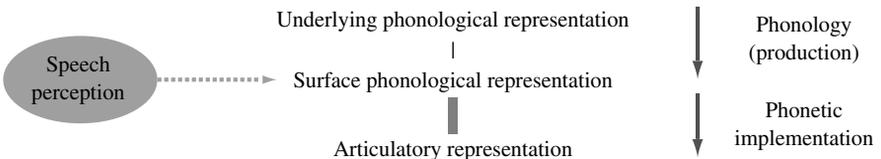


Figure 8.3 Three representations and the phonetics-phonology interface (thick line) as assumed in some phonological studies, with an indirect influence of speech perception (grey oval) on phonology. Processing is restricted to the production direction (arrows on the right)

differentiation between surface and articulatory form in Figure 8.3 does not hold for all of the aforementioned studies, since most of them belong to the minimal-functionalist approach elaborated in Section 3.1, and therefore collapse surface phonological and phonetic representations.

In contrast to the model given in Figure 8.3, one could consider speech perception as the mapping of a phonetic representation onto a surface phonological representation, thus approximately the reverse of phonetic implementation. If we include the perception/recognition direction of processing in our linguistic model (as forwarded in linguistics by, for example, Pierrehumbert, 1990), then the model looks as in Figure 8.4. In such a model, the phonetics-phonology interface is the mapping of surface onto phonetic form in the production direction and the reverse in the comprehension direction.

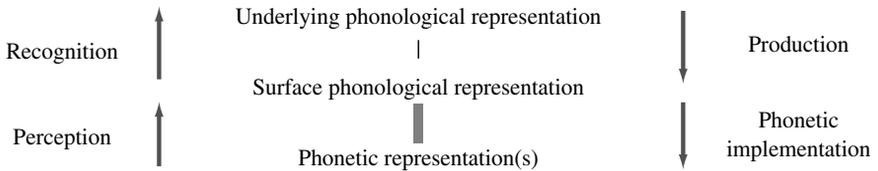


Figure 8.4 Two phonological representations and at least one phonetic representation, with two directions of processing (arrows left and right of the representations). The phonetics-phonology interface (bold line) depends on the number and ordering of the phonetic form(s).

In Figure 8.4, we see the comprehension process on the left and the production process on the right (as proposed in psycholinguistics by Cutler and Norris, 1979; McQueen and Cutler, 1997 and in linguistics by Boersma, 2007). In the present article, we will term the mapping of phonetic onto surface form in the comprehension direction ‘perception’, and the mapping of surface onto underlying form in the same direction ‘(phonological) recognition’. In the production direction, the mapping of underlying to surface form is ‘(phonological) production’, and the mapping of surface to phonetic form ‘phonetic implementation’ (following Boersma, 2007).

One phonetic representation does not seem to be sufficient in the model of Figure 8.4. In speech comprehension, listeners have an auditory form as input, whereas in speech production, speakers produce an articulatory form as output. It is therefore reasonable to assume that we have both an articulatory and an auditory phonetic representation. These two phonetic forms can be connected to the surface form in three different ways. First, both auditory form and articulatory form can be connected independently to the surface representation. Second, only the articulatory form can be connected to the surface representation, and the auditory form is then attached to the articulatory form.

And third, only the auditory form could be connected to the surface form, and the articulatory form is then connected to the auditory form. See Figure 8.5 for an illustration of these three possibilities.

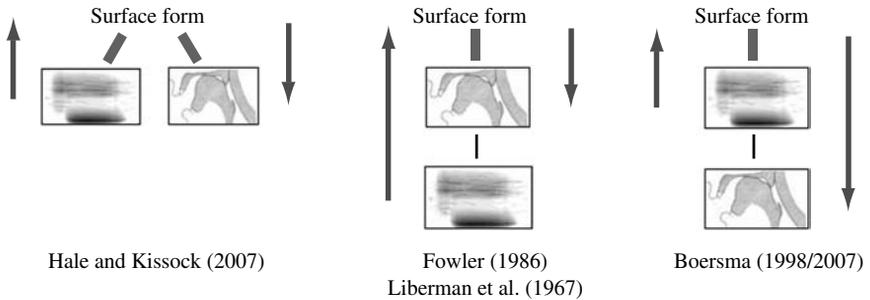


Figure 8.5 The three possible connections between the phonetic forms and the surface representation. The auditory representation is given as a cochleagram (of the German sound [ki], recorded and transformed with Praat: Boersma and Weenink, 2009), the articulatory representation is given as an x-ray tracing (of German /k/, from Wängler, 1958). The phonetics-phonology interface is indicated by a thick line. The directions of processing are indicated by the arrows; an upward arrow for perception, a downward arrow for phonetic implementation.

All three possibilities have been proposed. Model one and three we encountered already in Section 3.2. Hale and Kissock (2007) attach both the auditory and the articulatory form to the surface form, as in Figure 8.5 on the left. In this model, the two directions of processing involve different mappings: in perception from auditory form to surface form, in phonetic implementation from surface form to articulatory form. Boersma (1998/2007) connects only the auditory form directly to the surface form, as in Figure 8.5 on the right: in perception, the listener maps the auditory form onto the surface form; in phonetic implementation, the speaker does the reverse (with the same connections) and additionally maps the auditory onto the articulatory form.⁶

Figure 8.5 in the middle is the reverse of Boersma's; here the articulatory form is attached to the surface representation, and this connection is forming the phonetics-phonology interface. The auditory form is connected to the surface form only indirectly, and plays only a small role in speech perception and none in articulation. This view has been forwarded by Fowler (1986) as *Direct Realist Theory of Speech Perception*, which assumes that in speech perception the speech signal is directly interpreted in terms of articulatory gestures. A similar approach can be found in the *Motor Theory of Speech Perception* by Liberman and colleagues (Liberman et al., 1967; Mattingly and Liberman, 1969; Liberman

and Mattingly, 1985). According to this theory, we can only perceive speech sounds with the help of the gestures that we use for articulation. The interest in this theory was renewed with the discovery of mirror neurons, specifically with studies showing that there is activation in the mouth area of the motor cortex when participants listened to speech sounds (e.g. Fadiga et al., 2002; Watkins et al., 2003). Both the Direct Realism Theory and the Motor Theory of Speech Perception are, as their names imply, models of speech perception, and do not account for phonological processes.⁷

We saw now that the shape of the phonetics-phonology interface depends on a number of theoretical decisions. First, we have to decide whether we make a strict distinction between phonetic and phonological representations, that is, whether we employ a modular view of phonetics and phonology. If not, the interface is difficult to define. If we assume modularity, a further question is whether we want to include both directions of processing, speech production and speech comprehension, in our linguistic model. If we restrict ourselves to the production direction, then the interface is the connection between phonological surface forms and their articulatory correlates. If we include speech comprehension, we have to further ask ourselves what phonetic form is connected to the phonological surface representation. Here the interface can take any of the three possibilities given in Figure 8.5. Independent of the last two questions, we have to decide in a modular theory whether the connection between phonological surface form and phonetic representation(s) is automatic or language-specific. If it is automatic, then the phonetics-phonology interface is the same for all languages, if it is language-specific, we have to define the mapping between phonological surface form and phonetic form for every language.

In the following section, we will have a closer look at one phonological model that formalizes the interface as the language-specific mapping between surface form and auditory form. We will investigate what its underlying assumptions predict for the phonetics-phonology interface, and how these predictions relate to experimental findings in phonetics, phonology and psycholinguistics.

4. An Explicit Model of the Phonetics-Phonology Interface

The *Bidirectional Phonology and Phonetics* model (Boersma, 2007, 2009, to appear; henceforth BiPhon) is a modular approach to phonetics and phonology which includes both the production and the comprehension direction. It assumes an arbitrary connection between phonemes and their phonetic form, as promoted by Keating (1988) and Pierrehumbert (1990: 377ff.). All connections between representations, both phonological and phonetic, are formalized in the same way in BiPhon, namely as ranked OT-constraints.⁸ The representations and the constraints for their evaluation are given in Figure 8.6.⁹

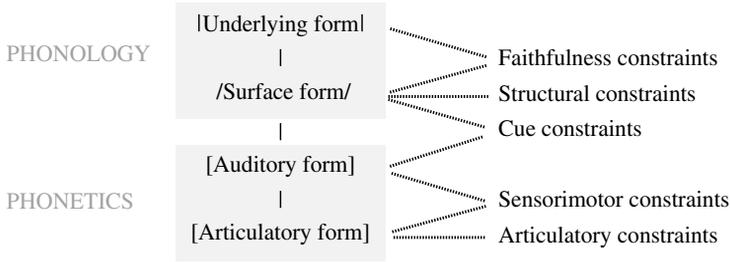


Figure 8.6 Representations and connections in the BiPhon model

The faithfulness constraints that map underlying to surface form and vice versa are well known from the OT-literature, where they are usually employed only in the production direction. The structural constraints replace partly the usually employed markedness constraints. In contrast to the latter, structural constraints only evaluate restrictions on the phonological surface form (such as *NoCODA, *COMPLEXONSET, TROCHEE etc.) but do not refer to any phonetic details (such as *RETROFLEX or MINDIST, recall the minimal functional approach described in Section 3.1). Restrictions on phonetic forms are formalized by cue constraints, sensorimotor constraints and articulatory constraints.

Cue constraints are responsible for the mapping of auditory cues onto phonological surface form. They constitute the phonetics-phonology interface. The cue constraint *[low F1]/a/, for instance, expresses the fact that the auditory cue of a low first formant should not be perceived as the vowel /a/.¹⁰ In the production direction, the same cue constraint is interpreted as the fact that the low vowel /a/ should not be realized with a low first formant. Examples for the language-specificity of cue constraints are given in Section 4.1. These cue constraints allow any arbitrary connection between a phonological category and an auditory dimension.

Sensorimotor constraints evaluate the mapping of articulation to auditory form. They ‘express the language user’s knowledge of the relation between articulation and sound; with them, the speaker knows how to articulate a given sound and can predict what a certain articulatory gesture will sound like’ (Boersma, 2009: 60). The sensorimotor constraints are usually not formalized, as they are not language-specific. Articulatory constraints regulate the restrictions on the articulatory form. An example is the constraint *DISTANCE(i-R), which stands for ‘avoid articulating sequences of high front vowels and retroflex sounds’ (Hamann, 2003: 180).

All constraints apply bidirectionally, that is, in both directions of processing. The strict division between phonology and phonetics is not only found in

the representations, but also in the constraints evaluating them, as we saw in the replacement of markedness constraints by structural constraints if they concern the phonological surface structure, and by cue or articulatory constraints if they concern phonetic representations (therefore the criticism on markedness constraints voiced by Haspelmath, 2006 and Hume, 2004/to appear does not apply).

4.1 The Language-Specific Phonetics-Phonology Interface

The BiPhon model assumes a language-specific mapping between auditory form and phonological surface representation, that is, a non-universal interface. An example from foreign-language perception will illustrate that this assumption and its formalization can account for observed phenomena. The example is the German perception of Dutch voiceless plosives, illustrated with the perception of the Dutch word *pitten* /.pɪ.tə./ 'to sleep'. This word is usually perceived by German listeners with no knowledge of Dutch as the German word *Bitte* /.bɪ.tə./ 'please'. The initial voiceless plosive is perceived as a voiced one because it lacks the aspiration that is typical for German initial voiceless plosives. Instead of the descriptive account given in the previous sentence, BiPhon allows to formalize this observation and to make testable predictions for other word forms and the perception and production of other languages by German native speakers.

In order to be able to account for the perception of /.pɪ.tə./ as /.bɪ.tə./, we have to specify the auditory realization of the Dutch word. The phonological form /.pɪ.tə./ has the auditory form transcribed here as [_Pt^ɪ_ə]. In this form, ' _ ' stands for the silence of the voiceless closure phase, 'P' and 'V' stand for the bilabial and alveolar voiceless release bursts, 't' and 'ə' stand for the respective formant values of the two Dutch vowels, and 't^ɪ' stands for the formant transitions into an alveolar stop (see Boersma and Hamann 2009b for similar transcriptions). This notation is employed to make the distinction between the abstract phonemes in /.pɪ.tə./ and the gradient auditory form that is associated with it explicit. How German listeners then perceive [_Pt^ɪ_ə] is determined by their language-specific cue constraints and the ranking of the constraints. These cue constraints connect the phonological categories that are employed in the language (in our restricted case /p/ and /b/) to the employed auditory dimension (here: aspiration noise [h], silence [_], and burst noise [P^b]). The cue constraints and the perception of [_Pt^ɪ_ə] by a German listener is shown in Tableau 8.1, where only those cue constraints that are relevant for the perception of the initial plosive are given.

Tableau 8.1 German perception of Dutch initial /p/, that is, the sound [p]

[Prt 'ə]	*[p]/.p/	*[p ^h]/.b/	*[p ^h]/.p/	*[p]/.b/
/.pr.tə/	*!			
☛ /.br.tə/				*

The cue constraint *[p] /p/ in Tableau 8.1 reads as ‘do not perceive the auditory form of a plain (i.e. unaspirated) labial plosive as an initial voiceless /p/’, the constraint *[p^h]/.b/ as ‘do not perceive the auditory form of an aspirated labial plosive as an initial voiced /b/’, and so forth. The constraints and their ranking, that is, the German perception grammar, mirror the fact that in German, plosive releases of initial voiceless labials with aspiration are usually perceived as /p/, and plosive releases of initial voiceless labials without aspiration are usually perceived as /b/.

The Dutch perception of the same auditory form results in a different output, as illustrated with Tableau 8.2. In the second and third cue constraint, ‘ ’ stands for a voiced closure phase, and ‘b’ for the bilabial voiced release bursts, the cue constraint *[b]/.p/ thus prohibits the perception of a voiced closure together with a voiced labial release as an initial voiceless /p/. The constraints and their ranking express the fact that in Dutch, plosive releases of initial voiceless labials without aspiration are usually perceived as /p/, and voiced closure phases with voiced labial releases as /b/.

Tableau 8.2 Dutch perception of Dutch initial /p/, that is, the sound [p]

[Prt 'ə]	*[p]/.b/	*[b]/.p/	*[b]/.b/	*[p]/.p/
☛ /.pr.tə/				*
/.br.tə/	*!			

We can see that with such formalization, the differences in phonetic details between languages can be made explicit and can account for differences in cross-language speech perception that have been observed in psycholinguistic studies (see, for example, Werker and Logan, 1985; Strange, 1995).¹¹

The same constraints and constraint ranking as in the perception processes are also used in the phonetic implementation of segments, where they can account for a foreign-language accent of an L2 speaker. If a German speaker with no knowledge of Dutch phonetics tries to produce the word *pitten*, she will use the same cue constraints and rankings as shown in Tableau 8.1, that is, her

native grammar. When used for phonetic implementation, the cue constraints have to be interpreted in the other direction, for instance, the constraint * $[_P]$ /p/ reads as ‘do not realize the surface form of an initial /p/ as the auditory form $[_P]$ ’, and so forth. The constraints and their ranking then formalize the fact that German initial /p/ is usually realized as voiceless plosive with aspiration noise, and initial /b/ as voiceless plosive without aspiration noise. In our example, we assume that the German speaker has created a surface form /pɪ.tən/ for *pitten*. The phonetic implementation process for German is shown in Tableau 8.3, for Dutch in Tableau 8.4, both focusing on the initial plosive.

Tableau 8.3 German phonetic implementation of an initial /p/

/pɪ.tən/	* $[_P]$ /p/	* $[_{ph}]$ /b/	* $[_{ph}]$ /p/	* $[_P]$ /b/
$[_{P}ɪt^{\text{h}}_tən]$	*!			
$[_{ph}ɪt^{\text{h}}_tən]$			*	

Tableau 8.4 Dutch phonetic implementation of an initial /p/

/pɪ.tən/	* $[_P]$ /b/	* $[_b]$ /p/	* $[_b]$ /b/	* $[_P]$ /p/
$[_bɪt^{\text{h}}_tən]$		*!		
$[_{P}ɪt^{\text{h}}_tən]$				*

The phonetic implementation in Tableau 8.3 thus accounts for the typical German aspiration of the voiceless, non-aspirated initial plosives in Dutch, cf. Tableau 8.4.

It could be argued that the differences we saw between German and Dutch phonetic implementation and perception do not lie in the mapping from auditory onto surface form and vice versa, but in a different surface (and underlying) representation. We could for instance assume that the Dutch /b/ – /p/ contrast is one represented with the feature $[\pm\text{voice}]$, whereas the German is one of the feature $[\pm\text{spread glottis}]$. This would allow for a universal phonetics-phonology interface, that is, a language-independent filling in of the auditory cues and articulatory gestures depending on the features used in the language (see Hale and Kissock’s 2007 approach in Section 3.2). For labial plosives, the feature $[\text{+voice}]$ would then translate into /b/, $[\text{–voice}]$ and $[\text{–spread glottis}]$ into /p/, and $[\text{+spread glottis}]$ into /p^h/. Two aspects oppose such a universal interpretation. First, there are small phonetic differences between languages with the same feature specification, for example, $[\text{+spread glottis}]$

means a stronger aspiration in Korean than in English (Kim, 1994; Kim et al., 2002), hence we would have to add some additional device to our theory to account for these differences. Second, there are phonological reasons independent of the phonetic realization for using the same phonological feature across languages. In our example, both Dutch and German /p/ are usually referred to by the phonological feature [-voice] (e.g. Booij, 1995, for Dutch; Wiese, 1996, for German), because in both languages /p/ groups with other voiceless obstruents. In German, /p/ forms a group with the other aspirated plosives /t/ and /k/ but also with the non-aspirated fricatives /f/, /s/ and /ʃ/ in the process of final devoicing, where voiced obstruents are devoiced in syllable-final position. The same process applies in Dutch, which leads phonologists to describe the process in both languages as a change in feature value from [+voice] to [-voice]. This identical formalization allows for an easier cross-linguistic comparison of the same or similar phonological processes. It is, however, a purely phonologically motivated use, to simplify cross-linguistic comparisons. It differs from the description with the phonetically more faithful features [±spread glottis] for German and [±voice] for Dutch.¹² In BiPhon, it is assumed that we do not have language-independent phonetic interpretability of features, for arguments given above. This allows us to use the same phonological features for the same processes.¹³

4.2 Bidirectionality of the Interface

In BiPhon, all constraints are used bidirectionally, that is, both in the comprehension and the production direction. We saw above that cue constraints form the phonetics-phonology interface in BiPhon. How do they behave in the two processing directions?

In perception, cue constraints interact with structural constraints. This means that speech perception is not only influenced by our native-language categories via the cue constraints (which map auditory information to phonemes or phonological features), but also by structural restrictions such as syllable structure and phonotactics. Psycholinguistic studies have shown that listeners cannot perceive a difference between words that deviate minimally from their native phonotactics and the corresponding correct words. This is illustrated with the well-known example of Japanese listeners who are not able to discriminate between [ebzo] and [ebuzo] (Dupoux et al., 1999). In the BiPhon model, this case of phonological knowledge influencing the perception is formalized as a structural constraint against coda consonants (*CODA, simplifying the Japanese facts slightly) being higher ranked than the cue constraint that militates against the interpretation of nothing in the auditory form as a high back unrounded vowel (*[]/u/).¹⁴

The formalization of perception as interaction of cue and structural constraints in BiPhon can furthermore account for the fact that loanword adaptation via perception often yields other outputs than we would expect based on native phonological processes. Boersma and Hamann (2009b) show that this is an automatic outcome of the model, and does not require loanword specific devices or constraint rankings as proposed in other studies (e.g. Rose and Demuth, 2006; Shinohara, 2004).

In the production direction, the use of cue constraints can explain why phonological surface structures can influence the phonetic realizations. This has been observed for instance in prosodic strengthening, where initial segments are increasingly 'strengthened' the higher they occur in the prosodic hierarchy. Strengthening can be realized in several ways such as lengthening, more amplitude and more linguo-palatal contact (see, for example, Fougeron, 2001; Wightman et al., 1992). In BiPhon, such strengthening is the result of position-dependent cue constraints, which are necessary anyway to account for, for example, syllable-dependent allophonic realizations.

In sum, the formalization of assumptions in BiPhon makes it possible to compare the outcome of this model with experimental data. Formalization can thus help us to have a grip on the phonetics-phonology interface and test the validity of the assumptions made in a model.

5. Conclusion: No Theory-Neutral Interface

The present article looked at several proposals on the phonetics-phonology interface. We saw that the definition of the interface depends strongly on our phonological assumptions. Do we want to give up on an abstract phonological form at all, as in Exemplar Theory? If not, do we want to employ a modular view of phonetics and phonology, that is, make a strict distinction between phonetic and phonological representations? If modularity is assumed, is perception part of the linguistic model or not? If perception is included, how is the auditory form integrated in the model? The scientist choosing a modular theory has to further decide whether the connection between phonological surface form and phonetic representation is automatic or language-specific. All of these choices result in a different phonetics-phonology interface.

In Section 4 we looked in detail at one proposal, namely the BiPhon model, with an explicit formalization of a modular theory that includes both directions of processing and works with language-specific mappings between phonological and phonetic form. We saw that this model makes predictions that can be checked against experimental data. Instead of arguing for a specific model, the aim of this section was to illustrate the need for explicit models and formalizations. A linguistic model that formalizes assumptions and therefore provides us

with testable predictions is preferable to approaches that simply describe observations. It is obvious that we need further models with explicit formalizations. Only alternative models based on different assumptions of phonology and its consequences for the phonetics-phonology interface can help us decide which phonological assumptions can best account for the observed data. Phonologists and phoneticians interested in the phonetics-phonology interface need to think in terms of models that can formalize the interface. These models, in turn, have to be informed again by experimental findings. As Fromkin (1975) has written:

[T]he study of phonology cannot be divorced from phonetics if phonologists aim to provide explanatory and predictive theories of sound systems, rather than merely descriptive theories. (p. 104)

6. Notes

1. For a detailed discussion of features and alternatives to the SPE feature set, the interested reader is referred to Botma et al., this volume.
2. This later point was already raised by Ladefoged (1972), who argued that different types of phonological descriptions require different types of features: a cross-linguistic comparison is best made with a restricted, unified set of categories (such as universal, phonetically-based features), whereas a complete description of a single language is best made with partly phonological features without phonetic correlates (to account for unnatural classes), and partly with features that have very detailed phonetic correlates (to account for small phonetic but non-distinct contrasts).
3. Cohn (2007) provides a more detailed discussion and counterarguments to this reduplication problem.
4. 'Minimal' is used to distinguish these approaches from other functional theories that also incorporate phonetic motivations of phonological processes into their formalism but follow strict modularity. For an example, see Section 4.
5. A similar point is made by Hale (2000: 243): 'It is common in phonological circles to use the symbols of the [...] IPA in a systematically ambiguous manner.'
6. Only in the phonetic implementation does Boersma's BiPhon model (2007) differ from his earlier work (1998), see Boersma and Hamann (2009b: 11f.) for details.
7. The primacy of the articulatory representation advocated in these two theories encounters two problems. First, infants can perceive speech sounds much earlier than they are able to produce them (e.g. Jusczyk, 1997). If speech perception depended on the availability of gestures, this would not be possible. Second, in speech production the target is assumed to be articulatory and not auditory, but bite-block experiments (Lindblom et al., 1979) showed that speakers can adjust to articulatory obstructions very quickly, which indicates that they most likely have an auditory rather than an articulatory target in their production process (Boersma, to appear: 22). These two points are also problematic for Hale and Kissock's proposal. The assumed equal status and independency of the auditory and the articulatory representation cannot account for the primacy of perception in the acquisition process. The findings in bite-block experiments are problematic since the articulatory form is assumed to be independent of the auditory form, thus a direct correction of articulation via the perceptual representation is not possible.

8. A formalization with weighted constraints, as in *Harmonic Grammar* (Smolensky and Legendre, 2006) is also possible, see, for example, Boersma and Escudero (2008) and Boersma and Hamann (2008). The connections could in principle also be formalized in a connectionist approach, see, for example, Plaut et al. (1996) on the connection between graphemes and phonemes, McClelland and Rumelhart (1981) and Rumelhart and McClelland (1986) on morphophonemic alternations, and McClelland and Elman (1986) on the modelling of speech perception.
9. The BiPhon model has been supplemented by a semantic module and respective representations, see Apoussidou (2007) and Boersma (to appear).
10. Cue constraints can also refer to other phonological categories such as features, syllables, feet etc. Syllable boundaries are for example, part of the cue constraints in section 4.1.
11. The German 'mis'-perception of [_P] as /b/ formalized in Tableau 8.1 could be argued to be due to a lexical effect, since German has the word *Bitte* /bɪ.tə./ 'please' but no word like *Pitte* with the surface form /pɪ.tə./. That the lexicon can influence the perception was shown by Ganong (1980) and Samuel (1987). However, an identification task with four native speakers of German without any knowledge of Dutch showed that the participants, who had to identify Dutch sequences of obstruent plus vowel /a/ as native German sequences, identified Dutch [_Pa] as German /ba/ in 42.98% of the cases (these results stem from the control group in Hamann and Sennema (2005) but are not reported in the article).
12. Here we see an illustration of Ladefoged's (1988) point that the use of features and their phonetic correlates, that is, the phonetics-phonology interface in Ladefoged's view, strongly depends on the aims of the phonological description (see endnote 2): if we aim at a phonological typological description, then we should use the same phonological features for the same/similar cross-linguistic processes, and if we aim at a universal phonetic interpretability of phonological features, we should use the same phonetic correlates for the same features. We cannot achieve both at the same time.
13. BiPhon is neutral with respect to the question whether phonological features are universal or not.
14. This constraint can only be low ranked in Japanese because the high back unrounded vowel /u/ is very often devoiced in Japanese and thus has often inaudible cues. Cue constraints are hence low ranked to explain the observed phenomena, but because the cues they refer to are not salient. High ranked cue constraints, on the other hand, mirror the fact that the cues they refer to are salient and important for the listeners of the respective language.

9 Syntax-Phonology Interface

*Anthi Revithiadou and
Vassilios Spyropoulos*

Chapter Overview

1. Introduction	225
2. The Syntax-Phonology Interface Defined: Basic Issues and Hallmark Theories	227
3. The Minimalist Program and the Syntax-Phonology Interface	241
4. Case Study: The Linearization of Subjects in Greek Clauses	245
5. Conclusions	251
6. Acknowledgements	252
7. Notes	252

1. Introduction

The flow and processing of information between the components of Grammar is a seminal topic in generative literature, the investigation of which has led to the development of elaborate theories about the architecture of grammar and the mode in which the interface among its components is performed. In this chapter, the focus is on a particular aspect of the interface and, more specifically, on the way syntax communicates with phonology. Although nowadays such intermodular relations are taken for granted and constitute an integral part of any theory that aspires at achieving a descriptive and explanatory efficacy, the syntax-phonology interface has not always been taken at face value.

One of the first attempts to ‘mix levels’ is encountered in Chomsky et al. (1956) where it is proposed that the distribution of stress in phonemic phrases in English (i.e. NP, VP and so on) is hierarchically organized and, furthermore, there is a relation between the phonological and higher levels of linguistic analysis (Chomsky et al., 1956: 78). For instance, the leftmost of two accented vowels is assigned primary stress prominence in phonemic phrases that contain an ‘internal juncture’ (i.e. compounds), for example, *lighthòuse*, whereas the

rightmost accent is promoted to primary stress in phrases which contain an 'external juncture', for example *light hóuse*. Kager and Zonneveld (1999), in a highly informative introduction on phrasal phonology, suggest that these ideas were supplemented 10 years later by an inspiring, yet not oft-cited paper by Bierwisch (1966), who explicitly argues that 'syntax feeds phonology' in an indirect way: the output of syntax is converted by means of a mapping procedure into the input of phonology. In the coming years, however, the issue of the interface was not fervently pursued.

The *Sound Pattern of English* (SPE) (Chomsky and Halle, 1968: 9) recognizes an output of the syntactic component, an input to the phonological component and readjustment rules whose main task is to convert the syntactic string into a form that can be read off and interpreted by phonology. Syntactic information is encoded in phonology by rules that insert *boundary* symbols, that is, segment-like elements, which are deprived of any phonetic manifestation, at the edges of syntactic constituents. For instance, the boundary '#' is used to indicate edges of major syntactic categories (N, V, A etc.) and phrasal categories (e.g. NP, VP etc.).¹ However, SPE has very little to contribute in the way these components interact. In a sentence like *This is the cat that caught the rat that stole the cheese*, the intonational breaks after *cat*, *rat* and *cheese* do not match the multiply embedded structure of the syntactic string. Although Chomsky and Halle (1968: 372) seemed to be willing to attribute the 'flatter' phonological structure to readjustment rules, they eventually accredited the mismatch to the interference of performance factors, which fall outside of Grammar (cf. Nespor and Vogel, 1986 for argumentation against this view). Much of the post-SPE era is devoted to defining the exact number of boundaries and their relative strength (see McCawley, 1968; Selkirk, 1972; Sag, 1974; Basbøll, 1975, 1981; and Scheer, 2006, 2008, in press a, for a recent reincarnation of the boundary theory). As new advances in syntactic theory delineated the future of generative grammar, the focus of attention gradually shifted from the typology of readjustment rules to issues pertaining to: (i) the direct or indirect reference relation between the relevant components of the grammar, (ii) the type of syntactic information phonology has access to, (iii) the exact nature of the syntax-phonology interaction and, especially, the question whether syntax feeds phonology (serialism) or the two modules are co-present and interact simultaneously (parallelism) and (iv) the procedures or the operations that take place at this interaction. As far as the latter issue is concerned, there is a growing interest in linearization, that is, the procedures that translate the hierarchically organized syntactic structure into the linearly ordered phonological structure and decide on which syntactic objects will be pronounced or silenced.

In the remainder of this chapter, we highlight the main developments in this area of research (Section 2) and then focus on the parallelism versus serialism question (Section 3) which is empirically explored with the help of a case study

from Greek (Section 4). More specifically, in Section 2 we review some early studies on the syntax-phonology interface that played a crucial role in the shaping of the early stages of the field. Then, we move on to examining the factors and developments that led to the *prosodic structure hypothesis* (Selkirk, 1981a, b, 1984b, 1986, 1995a; Nespor and Vogel, 1982, 1986) and the introduction of a theory of phonological representations, namely *Prosodic Phonology*, as the counterpart of the *direct access hypothesis*, which maintains that phonology has immediate access to syntactic word and phrase structure (Kaisse, 1983, 1985 and following; Odden, 1987, 1990, among others). In Section 3, emphasis is on the serialism versus parallelism debate. In light of recent developments in syntactic theory, it is argued that phonology has a limited effect on syntax, in the sense that it can only endorse certain selections which become available to phonology during the linearization of constituents in a syntactic string. This particular view of serialism is empirically substantiated in Section 4 with evidence from Greek word order in subject constructions. Section 5 concludes the chapter.

2. The Syntax-Phonology Interface Defined: Basic Issues and Hallmark Theories

2.1 The Direct Access Versus the Prosodic Structure Debate²

Chomsky's (1970) X-bar theory of phrase structure had a profound and constructive impact on the way researchers viewed the syntax-phonology interface. Selkirk (1972, 1974) phrases the sandhi rule of French *liaison* in terms of X-bar notation. The rule concerned deletes a word-final obstruent. In so-called *liaison contexts* the rule fails to operate, that is, the final consonant is preserved, when the following word is vowel-initial (1a); with consonant initial words, the consonant is omitted (indicated with a <C>) (1b).

- (1) a. des ennuis
 'troubles'
 Paul nous appelle
 'Paul is calling us'
- b. de<s> problèmes
 'problems'
 Paul nou<s> repousse
 'Paul rejects us'

Notice that the terminal strings where the rule applies contain a single word boundary ('#') *des # ennuis, des # problèmes, Paul # nous # appelle* and so on,

since a determiner or a clitic pronoun precedes a major category element, that is, a noun or a verb. In short, the rule of final obstruent deletion fails to apply when the consonant at issue is followed by a # boundary and a vowel-initial word. Interestingly, the rule is also blocked in certain syntactic environments such as, for instance, between a subject noun phrase and a verb (2a-b) and between the verb and the object noun phrase (2c-d).

- (2) a. Les garçon<s> enragent (Selkirk, 1974: 580)
'The boys are getting mad'
b. Les garçon<s> pleurent
'The boys are crying'
c. Les immigré<s> envoyaien<t> des lettre<s> à leurs familles
'The immigrants were sending letters to their families'
d. L' immigré envoyai<t> un paquet<t> à sa famille
'The immigrant was sending a package to his family'

The contexts of liaison vary slightly depending on the style of speech. The statement in (3) holds true in an elevated style of French.

- (3) A head Noun, Verb or Adjective which is inflected may be in a liaison context with the word that follows, if that word is in its Complement (= sister to N, V, A). (Selkirk, 1974: 581)

Based on the above examples, we conclude that liaison applies between the elements of a major category. An inflected noun or verb will not be in a liaison context with its non-complement, that is, any element outside its phrase, as shown in (4). This entails that a readjustment rule, called X-Comp, applies to erase one of the two ## boundaries that naturally separate the major category items in the relevant string. Furthermore, the example set in (4c) and (5c) demonstrates that such sandhi rules have a key-role in solving ambiguities; therefore, they should be considered an integral part of the grammar of the language and not a part of performance (cf. Chomsky and Halle, 1968).

- (4) a. des endroits obscurs (Selkirk, 1974: 583)
'dark places'
b. des promenades en voiture
'rides in the car'
c. un marchand de draps anglais
'a merchant of English sheets'
- (5) a. Donnez ces lunette<s> à Marcel (Selkirk, 1974: 584)
'Give these glasses to Marcel'

- b. J'ai trouvé mes lunettes<s> à la mairie
'I found my glasses at the town hall'
- c. un marchand de draps<s> anglais (Selkirk, 1974: 583)
'an English merchant of sheets'

The importance of Selkirk's work lies in the fact that X-Comp refers *directly* to syntactic information and hence lies at the core of the syntax-phonology interface. Napoli and Nespor (1979) take a more radical approach to the interface by eliminating readjustment rules and boundaries altogether and making phonology explicitly dependent on syntax. In particular, they propose that the rule of *raddoppiamento sintattico* (RS), one of the trademarks of the Sicilian and Tuscan varieties of Italian (6), applies between two elements under the proviso that the leftmost one is a left-branch constituent, as stated in (7).

- (6) a. Fa [k:]aldo (/fá káldo/ →[fá k:áldo]) (Napoli and Nespor, 1979: 812)
'It's hot'
- b. Ho visto tre [k:]ani (Napoli and Nespor, 1979: 825)
'I saw three dogs'
- c. Mario ha [f:]atto tutto
'Mario did everything'
- d. Ho visto tre [g:]randi cani
'I saw three big dogs'
- e. Ha [s:]empre parlato bene di te
'He's always spoken well to you'

- (7) *Left Branch Condition*: RS can apply between a word *a* and a following word *b*, where *a* is immediately dominated by the preterminal category symbol *A*, and *b* is dominated (not necessarily immediately) by the category symbol *B*, only if *A* is a left branch of the first node that dominates both *A* and *B*.

Given the condition in (7) and the language-specific order and number of the complements of X in (8), RS is expected to apply between a Specifier and a following word, that is, an X (6b-c) or a complement of X (6d-e). In the examples (6b-c), for instance, RS applies between the Spec and the N and V, respectively, of the relevant major category (X'') (i.e. NP and VP, respectively).

- (8) [_{X'} Specifier [_X (complement) X (complement) (complement) . . .]]

The proposed analysis acknowledges that phonology and (surface) syntactic structure are closely connected and, furthermore, syntactic notions such as 'X-bar constituent', 'major category' and 'branch' are exploited in full by

phonology. In subsequent years, this line of research was further extended by proponents of what is broadly known as the *direct access hypothesis*, most notably expressed in the works of Kaisse (1983, 1985, 1990); Kaisse and Zwicky (1987); Odden (1987, 1990), among others. Kaisse (1985), on the basis of cross-linguistic evidence, proposes a direct reference model which distinguishes between post-lexical rules which are sensitive to morphosyntactic information (e.g. rules of external sandhi) and those that are not (e.g. fast speech rules).³ For instance, in Athenian Greek the rule of *unrounded first-vowel deletion* applies between a noun and its preceding material (9a), and between a verb and the (following) complements (9b) (Kaisse, 1985: 125).

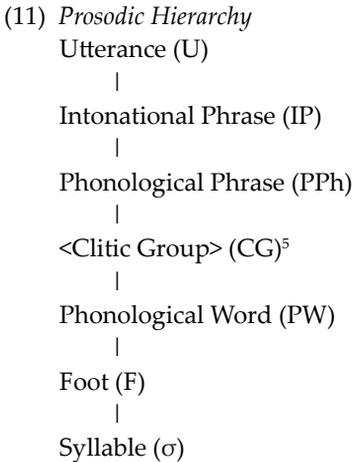
- (9) a. /ta meɣála eláfia/ [ta meɣál eláfça] (Kaisse, 1985: 116)
 the big-PL deer-PL
 'The big deers'
 (cf. *[to meɣál eláfi] < /to meɣálo eláfi/ 'the big deer')
- b. /xtípise aýória/ [xtípis aýórja]
 hit-PAST.3SG boy-ACC.PL
 '(S)he hit boys'
 (cf. *[alaz aýórja] < /alázo aýória/ 'I change boys')

Interestingly, this rule does not apply in all environments where its structural description is met. For instance, in *NP-Subject V* strings, its effects are superseded by another rule which blindly deletes the last vowel of the first element in the string. The main point of this approach, therefore, is that phonological rules incorporate syntactic information in their structural description. In our case, the rule that deletes unrounded vowels operates in certain syntactic strings. The natural question of course is what type of syntactic information phonology may have access to. According to Kaisse (1985 and following), *c-command* is a syntactic relation which is visible to phonology as shown in (10):

- (10) *Domain c-command*: In the structure $[x^{\max} \dots \alpha \dots]$, X^{\max} is defined as the domain of α . Then α c-commands any β in its domain.
 (Kaisse, 1985: 159)

Domain c-command implies that for a rule to apply to a string $\alpha\beta$, α must domain c-command β . In other words, c-command can operate from α into maximal projections but not to any constituent out of them. Thus, in (9b) the verb *xtípise* c-commands the noun *aýória*, in (9a) the head of the NP, that is, *eláfia*, c-commands the determiner *ta* and everything included within the AP, that is, *meɣála*.

At the same time, new advances in metrical phonology which assign a tree-like structure to phonological representations (Lieberman, 1975; Liberman and Prince, 1977) led to the introduction of the notion of hierarchy in phonology. Selkirk (1980a [1978], 1980b, 1981a) takes this idea one step further and argues against a linear arrangement of segments and boundaries and in favor of a ‘suprasegmental, hierarchically arranged organization’ of the utterance (Selkirk, 1981a: 111).⁴ More specifically, she argues that phonological representations are arranged into binary branching trees (with strong/weak labelled nodes) and, more importantly, form specific prosodic categories. These prosodic units constitute delimiting domains for the application of phonological rules and phonotactic conditions. That is, phonology cannot directly access syntax nor can it read off syntactic information straight from arboreal structures. Phonological rules can only refer to prosodic constituents such as the syllable, the foot, the phonological word, the phonological phrase, the intonational phrase and the utterance. The hierarchical constellation in (11), known in the literature as the *Prosodic Hierarchy* (PH) (Selkirk, 1980a [1978] and following; Nespor and Vogel, 1982, 1986; Hayes, 1989b [1984]), signals the birth of *Prosodic Phonology*.



The major prosodic units that mediate the syntax-phonology interface are the *prosodic word* (aka *phonological word*, PW) and the *phonological phrase* (PPh). Two types of PPhs are distinguished in the literature: (i) the *minor phrase* often referred to as the *accentual phrase* and (ii) the *major phrase* (or *intermediate phrase*) (cf. Beckman and Pierrehumbert, 1986; Pierrehumbert and Beckman, 1988; Selkirk and Shen, 1990; Selkirk, 2000, among others). It is the job of a mapping mechanism to translate syntactic structure into prosodic structure, that is, a PW and a PPh. Selkirk (1981a) proposes that the mapping rule in

English, for example, operates on stems to construct feet and assemble them into a right-branching construction. The representation in Figure 9.1 constitutes a PW. Similarly, a mapping rule singles out phrasal heads in a syntactic string and organizes the material before the head (e.g. Det, Prep, Comp, Aux etc.) as well as the head itself into a right-branching structure which constitutes a PPh, as shown in Figure 9.2. That is, the specifier of a syntactic phrase joins with the head of the phrase into a PPh and a non-lexical item joins with its sister constituent. Within a PPh, the rightmost element is labelled *s* and its sister node is labelled *w*.

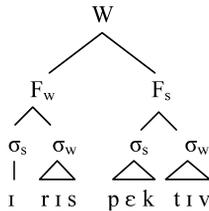


Figure 9.1 Prosodic word: ‘irrespective’ (Selkirk, 1981a: 125–127)

[_S [_{NP} Det Adj N] [_{VP} V_{Aux} V_{Aux} Adv V [_{PP} P [_{NP} N]]]

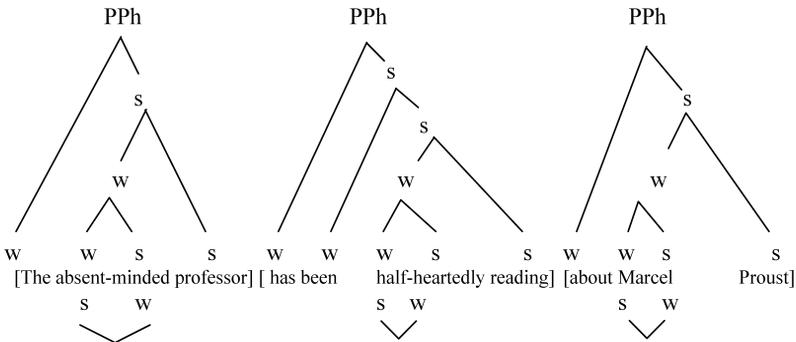


Figure 9.2 Phonological phrases: ‘The absent-minded professor has been half-heartedly reading about Marcel Proust’ (slightly modified from Selkirk, 1981a)

As shown in Figure 9.2, the VP splits into two PPhs thus indicating the lack of isomorphism between the syntactic and phonological component. Evidence for the existence of PPh in English can be drawn from the *rhythm rule* of English (Lieberman and Prince, 1977), which reverses the *ws* pattern of *absent-minded*, for instance, into *sw* so that a stress clash, that is, a sequence of two *s*'s in the same tree, can be avoided. The rule is enforced within the PPh domain, as suggested

by the tolerance of two adjacent prominences in sentences like (12) where the adjacent *s*-nodes belong to different PPhs.

- (12) w s s
 [Marcel]_{PPh} [read]_{PPh} [about the professor]_{PPh}

The classical Prosodic Phonology theory (Selkirk, 1981a and b, 1984b, 1986; Nespor and Vogel, 1982, 1983, 1986) assumes the organization of grammar in (13). Rules that undertake the mapping of morphosyntactic structure into a PW and PPh are part of the *post-lexical* phonology. Selkirk (1984b: 82) dubs this a *syntax-first* model.

- (13) syntactic surface structure → mapping rules →
 prosodic representations → phonological rules →
 phonetic representations

To sum up, the *prosodic structure hypothesis* holds that ‘the phonological representation of a sentence is organized into a prosodic constituent structure which is independent of, but related to, the surface syntactic structure of a sentence’ (Selkirk, 2002: 15408, column A and references cited therein). The existence of such a mapping device clearly entails that syntactic constituents and relations (e.g. *c-command*) and ‘empty’ syntactic entities (e.g. *pros*, *wh*-traces etc.) are not visible and hence accessible by rules of phonology proper (i.e. rules inducing changes in the phonological pattern of a string of elements). Phonology communicates with syntax via the prototypically interface sites, namely the PW and the PPh.

2.2 Argumentation for Prosodic Constituency: (Non-)isomorphism

The pivotal argument in favour of prosodic constituency is that prosodic structure may reflect syntactic structure but need not be isomorphic to it. We will further exemplify the lack of isomorphism in the syntax-phonology interface at the PW- and PPh-level with the help of two examples from Greek and Italian.

In Greek, weak object pronouns encliticize to imperative forms and gerunds (14) (Revithiadou and Spyropoulos, 2008, and references cited therein).

- (14) *Greek object clitics*
 a. tráva to / travóⁿdas to
 pull-2.SG.IMP CL.3-N.SG.ACC / pull-GER CL.3-N.SG.ACC
 ‘Pull it!’/ ‘While pulling it’

- b. *đjavasé* to / *đjavázonⁿdás* to < *đjavase*, *đjavázonⁿdas*
 read-2.SG.IMP CL.3-N.SG.ACC / read-GER CL.3-N.SG.ACC
 'Read it!' / 'While reading it'

The verb form and the clitic constitute a single prosodic unit, namely a PW. This is primarily evidenced by the development of rhythmic stress, as shown in (14b). The addition of the clitic causes the form to be further footed so that the three-syllable restriction imposed by the language can be salvaged: (*đjáva*) (*sé to*). It is obvious that in this case the PW exceeds the boundaries of the lexical word since it incorporates more material into its domain. This is a typical case of non-isomorphism.

Another informative example of non-isomorphism is taken from Nespor and Vogel (1986), who re-analyse the RS data in light of prosodic phonology. More specifically, RS is argued to apply within the PPh but it is blocked between elements that belong to different PPhs.

- (15) a. [Ho [v:]isto]_{PPh} [tre [k:]olibri]_{PPh} [molto scuri]_{PPh}
 'I saw three very dark hummingbirds'
 b. [Ho [v:]isto]_{PPh} [tre [k:]ani]_{PPh}
 'I saw three dogs'
 c. [Avrá [t:]rovato]_{PPh} [il pescecane]_{PPh}
 'He must have found the shark'

Nespor and Vogel (1986) explain that the string in (15c), for instance, which is organized into two PPhs, represents a single syntactic constituent, namely, a VP (or a V'' in the X-bar theory jargon). Similarly, although the chunk *tre colibri molto scuri* in (15a) belongs to the same syntactic constituent, namely, NP (=N''), it is parsed into two different PPhs, as evidenced by the failure of RS to apply between *colibri* and *molto*. In other words, a syntactic phrase is split into two autonomous PPhs.

Since phonological rules have a piece of structure as their domain of application in the realm of which phonotactic restrictions and stress prominence relations may also be enforced, it is important for a theory such as prosodic phonology to determine the diagnostic criteria that establish the notion of *constituent* in phonology (Nespor and Vogel, 1986: 58ff.). Circularity in the application of a phonological rule as a criterion for prosodic constituency is avoided only if a number of phenomena cluster together in order for a particular string of elements to be matched with a particular prosodic unit. A wide range of diagnostic criteria have been put forward to identify the domain of the PW (see Revithiadou, to appear, for an overview) and the PPh

(see Selkirk, 2002, for an overview). Some of the well established ones for the identification of the PW and the PPh are: various segmental rules (e.g. harmony processes, assimilation, dissimilation, various types of deletion and epenthesis rules etc.), the distribution of word stress and lexical tone(s), the distribution of phrasal boundary tones, (re-) syllabification, phonotactic conditions and so on.

Significantly, Selkirk (1995a) and Vogel (2009) also point out that prosodic constituents such as the PW and the PPh, which derive from the interface of phonology with the other components of the grammar, are also identified by the mapping relations. Thus, an elaborate and careful analysis of the morphosyntax of a language may also prove valuable in establishing prosodic constituency. As a result, the focus of research has gradually shifted towards addressing important questions regarding the exact nature of the mapping algorithm, the way it is performed and, in particular, the direction of the intermodular interaction. For instance, does syntax feed phonology (serial model) or the two components operate in parallel and mutually influence each other (bidirectional model)? These questions are more articulately addressed in the following sub-section.

2.3 Mapping Mechanisms and Mode of Intermodular Interaction

Constructing mapping algorithms has not been – and still is not – an easy task because it primarily depends on defining the nature of the syntactic information that is relevant to prosodic phrasing. There are two major standpoints in the phrasal phonology literature regarding the mapping mechanism. First, there is the *relation-based* (RB) model, advocated mainly in the works of Nespor and Vogel (1982, 1983, 1986) and Hayes (1989b [1984]), which assigns a key-role to the tree geometry and, especially, the notion of recursion as well as the structural relations between syntactic constituents (e.g. head-complement, argument-adjunct). Second, there is the *end-based* (EB) model, originally proposed by Selkirk (1986 and following) and later further developed by Selkirk and Tateishi (1988, 1991), Selkirk and Shen (1990) and Truckenbrodt (1995, 1999). The EB theory puts emphasis on node labels and asserts that the mapping accesses surface syntactic structure only and, more specifically, edges of lexical categories. Although these approaches to mapping do not necessarily stand in a rivalry relationship and may in certain ways even complement each other (cf. Chen, 1990), it has been claimed that the latter is theoretically more restrictive and empirically more justified than the former (cf. Selkirk, 1986; Inkelas and Zec, 1990).⁶ The common denominator in both approaches is that they take a ‘syntax-first’ stance in the mode of interaction between the relevant components: The output of syntax is inserted into phonology in order to be

assigned prosodic structure. Below we lay out the main principles of each mapping algorithm.

Nespor and Vogel (1986: 168) propose the mapping rules in (16) as part of a universal mechanism of PPh construction.

(16) *PPh formation*

- a. The domain of *PPh* consists of *C* (Clitic Group) which contains a lexical head (*X*) and all *C*s on its non-recursive side up to the *C* that contains another head outside of the maximal projection of *X*.
- b. Join into an *n*-ary branching *PPh* all *C*s included in a string delimited by the definition of the domain of *PPh*.

Recursion is a central notion in syntax since it is a parametric choice of an individual language whether its complements and specifiers will appear at the left or right edge of the head of a syntactic phrase. The by now familiar example from Tuscan, given in (15a) and repeated in (17), is divided into three PPhs. According to (16), each lexical head (i.e. *visto*, *colibri* and *scuri*) will form a PPh with all elements on its non-recursive side. Since Italian is syntactically right branching, the syntactic head will prosodify into a single PPh with all elements at its left.⁷

- (17) a. [_{VP} Ho visto [_{NP} tre colibri [_{AP} molto scuri]]]
 b. [Ho [**v**:]isto]_{PPh} [tre [**k**:]olibri]_{PPh} [molto scuri]_{PPh}
 'I saw three very dark hummingbirds'

The new element that this approach brings to the interface debate compared, say, to the direct interface analysis proposed a few years earlier by Napoli and Nespor (1979) is the notion of optional restructuring as shown in (18).

(18) *PPh-restructuring* (optional)

A non-branching PPh which is the first complement of *X* on its recursive side is joined into the PPh that contains *X*.

That is, if a right-hand complement is not a branched PPh, it may optionally join to the preceding PPh, as indicated in the following example.

- (19) a. [I caribú]_N_{PPh} [nani]_A_{PPh} sono estinti
 b. [I caribu]_N **n**:ani]_A_{PPh} sono estinti *restructuring*
 'Dwarf caribous are extinct'

The optional application of RS in (19b), in contrast to the obligatory application of the rule in other examples (e.g. (17b)), is predicted to occur in certain types of

PPhs, namely those that are the by-product of restructuring between a head and its right-hand material due to the lack of branchingness of the latter (e.g. *nani* in (19)). Had it been a branched structure, PPh-restructuring would not have applied. The condition under which restructuring applies suggests that the length (and not just the syntactic status) of a constituent may be relevant for the application of a phonological rule and phonology, in general.

Other RB approaches to the interface evoke various sorts of relations holding between syntactic constituents that belong to the same PPh. Hayes (1989b [1984]) calls upon complement and head adjacency for PPh formation in ChiMwiini; whereas, in his analysis of Xiamen Chinese, Chen (1990) employs RB rules which rely on adjacent constituents involved in head-argument and head-adjunct relations. Several RB models need to invoke an extra proviso on adjacency, which is costly and not advantageous, in terms of descriptive adequacy, compared to the EB-model, discussed in the ensuing paragraphs.

The EB theory of mapping is characterized by its limited sensitivity to syntactic structure. Selkirk (1986), following Chen (1985), proposes the EB mapping theory which operates on the edges of X-bar constituents. The basic idea is that a prosodic constituent is demarcated by the right or the left edge of selected syntactic constituents (i.e. X^0 , X' , X'') as shown in (20) (Selkirk and Shen, 1990: 319).

(20) *The syntax-phonology mapping*

For each category C^n of the prosodic structure of a language there is a two-part parameter of the form

$$C^n: \{Right/Left; X^m\}$$

where X^m is a category type in the X-bar hierarchy.

An advantage of this model is that the mapping algorithm is cross-categorical: the end-rules can apply at some level of the X-bar hierarchy in order to form the appropriate prosodic domains. For instance, they apply at X^0 to derive PWs and at X' or X'' to derive (minor and major) PPhs. Differences in mapping between languages are due to parametric variation with respect to the X-bar level and the relevant edges. For instance, in the abstract examples in (21), the edge of the X^0 (=word) or the XP which will serve as the beginning or the end point of the PW and PPh domain, respectively, is decided on a language-particular basis.

$$(21) /X^0 \ y \ X^0/ \qquad \qquad \qquad /XP \ y \ XP/ \qquad \qquad \qquad \text{where } y: \text{det, compl, conj etc.}$$

$$\text{a. } \begin{matrix} \text{PW}[X^0 \ y \ \text{PW}[X^0} & \text{end: left} & \text{PPh}[XP \ y \ \text{PPh}[XP} \\ \rightarrow [X^0 \ y]_{\text{PW}}[X^0]_{\text{PW}} & & [XP \ y]_{\text{PPh}} [XP]_{\text{PPh}} \end{matrix}$$

$$\text{b. } \begin{matrix} X^0]_{\text{PW}} \ y \ X^0]_{\text{PW}} & \text{end: right} & XP]_{\text{PPh}} \ y \ XP]_{\text{PPh}} \\ \rightarrow [X^0]_{\text{PW}} [y \ X^0]_{\text{PW}} & & [XP]_{\text{PPh}} [y \ XP]_{\text{PPh}} \end{matrix}$$

In the above examples, function elements will be prosodified together with lexical categories. Selkirk (1995a) claims that there is a systematic bias towards lexical categories, which is also reflected in the process of language acquisition. Thus, the distinction between lexical and functional categories is crucial in the mapping of syntactic structure to prosodic structure and hence relevant to the mapping rules (cf. *Lexical Category Condition* (LCC); Truckenbrodt, 1999: 226).

Let us illustrate the EB mapping with the help of an actual example. In Shanghai Chinese, the Major Phrase constitutes the domain for the application of Post-Focus Tonal Deletion (Selkirk and Shen, 1990). The Major Phrase is constructed by the rule: {Left, Lex^{max}}. That is, in the sentence in Figure 9.3, a MaP break will fall at the left edge of the: (a) PP, (b) the highest pre-verbal NP, 'oN 'geq ?njiaw 'geq, (c) the next lower NP, ?njiaw 'geq and (d) the post-verbal NP, resulting in a prosodic structure that is highly non-isomorphic to syntax:

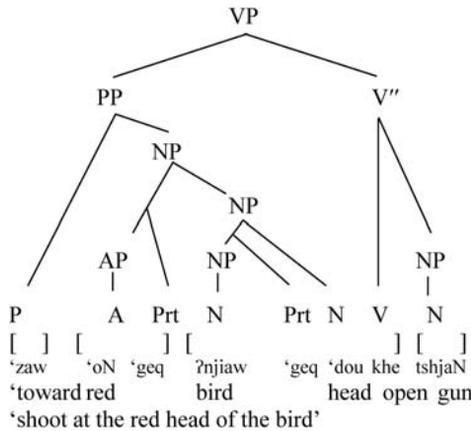


Figure 9.3 MaP construction according to EB mapping (Selkirk and Shen, 1990: 329–330)

Notice that the pre-final MaP extends from the middle of the NP-object of the pre-verbal PP to the middle of the V''. Focusing of 'bird' would cause deletion of tone in the following elements of the same MaP but would not force tone deletion in the noun 'gun' because the latter belongs to a different MaP.

In Optimality Theory (OT) (Prince and Smolensky, 2004 [1993]), the edge-based rules translate into the constraint-format of McCarthy and Prince's (1993a) *Generalized Alignment*. In the Optimality theoretic spirit, Selkirk (1995a) proposes

the constraint in (22a) on edge-alignment of syntactic phrases with phonological phrases. In addition, Truckenbrodt (1995, 1999) provides compelling arguments for the necessity of including the cohesional constraint WRAP-XP in (22b), in the family of interface constraints. In many languages, a major syntactic phrase preserves its integrity and is mapped into a single PPh. In accordance with the LCC, the constraint penalizes separate PPh of lexical projections but, interestingly, permits the split up of functional ones.

(22) *edge-alignment constraint*

- a. ALIGN-XP: Align (XP, L/R; PPh, L/R)
For each XP, there is a PPh such that the left (/right) edge of XP coincides with the left (/right) edge of a PPh.
- b. WRAP-XP
Each XP is contained in a PPh.

The abstract example in (23) illustrates the effects of ALIGN-XP and WRAP-XP. Differences in p-phrasing across languages result from different rankings of the relevant constraints.

- | | | |
|------|--|---------------------------------------|
| (23) | [_{VP} V NP PP] | <i>syntactic string</i> |
| a. | [] _{PPh} [] _{PPh} [] _{PPh} | PPh due to high-ranking of ALIGN-XP,L |
| b. | [] _{PPh} [] _{PPh} | PPh due to high-ranking of ALIGN-XP,R |
| c. | [] _{PPh} | PPh due to high-ranking of WRAP-XP |

In a brief comparison of the two mapping algorithms, it is evident that EB is more restrictive than RB in the sense that only the edges of lexical syntactic entities are visible to phonology. Functional information and other type of structural relations that hold among the elements of a given chunk do not constitute relevant information for phonology. RB seems to gain empirical advantage by allowing some leeway in PPh-restructuring: non-branching elements may rephrase by joining the PPhs of the neighboring heads. In fact, it turns out that branchingness may play an even more important role than it has so far been assumed.⁸ Ghini in his 1993 analysis of phrasing in Italian shows that, if syntactic branchingness is re-analysed as phonological branchingness, the mapping algorithm enjoys greater empirical support. Furthermore, Selkirk (2000) underlines the importance of constraints on the minimum and maximum size of prosodic constituents as the driving force for eurhythmic phrasing patterns typically associated with fast or informal speech styles. She proposes two size constraints on PPhs, stated in (24), which assess the well-formedness of a constituent of a particular level of prosodic structure Cⁱ in terms of the number of constituents of a particular level Cⁱ⁻¹ that it contains (Selkirk, 2000: 244).

- (24) *Binarity constraints* (Selkirk, 2000, based on Itô and Mester, 1992, 1995)
- a. $\text{BINARITY}^{\text{min}}$
A PPh must consist of at least two PWs.
 - b. $\text{BINARITY}^{\text{max}}$
A PPh must consist of at most two PWs.

With the constraints in (24) ranked above ALIGN-R , which is ranked higher than ALIGN-L in Greek phrasing, this will yield the output in (25b) for an input string such as *to fos ðíni isxí sti mixaní* ‘the light gives power to the engine’ in Greek. The application (or blocking) of *s*-voicing before a voiced fricative or nasal and vowel degemination are the diagnostics for each phrasing option.

- (25) $[\text{I}_{\text{IndP}} [\text{N}_{\text{P}} \text{Det N}] [\text{I}_{\text{Ind}} \text{V}] [\text{V}_{\text{P}} \text{t}_{\text{V}} [\text{N}_{\text{P}} \text{N}] [\text{P}_{\text{P}} \text{P NP}]]]$
- | | | | | | | |
|---------------------------------------|-----------|----------|-----------|--------|------------|--|
| to | fós | ðíni | isxí | sti | mixaní | |
| the | light-NOM | give-3SG | power-ACC | to-the | engine-ACC | |
| ‘The light gives power to the engine’ | | | | | | |
- a. $[\text{to fós}]_{\text{PPh}} [\text{ðín} \emptyset \text{ isxí}]_{\text{PPh}} [\text{sti mixaní}]_{\text{PPh}}$ *EB mapping*
 - b. $[\text{to fóz ðíni}]_{\text{PPh}} [\text{isxí sti mixaní}]_{\text{PPh}}$ *binarity effects*

To sum up the discussion so far, two approaches to mapping have been proposed in the literature: The EB-model, which is blind to syntactic relations since it aims at assigning a PPh-boundary to the left/right edge of an XP, and the RB-model, which is sensitive to the architecture of the syntactic tree and the relations that hold among its constituents. Both are enriched with an optional restructuring mechanism which takes the form of either an optional rule or, in OT terms, a(n) (alternative) ranking in which binarity constraints are ranked higher than the relevant interface constraints.

Zec and Inkelas (1990) take the branchingness argument one step further and propose that it plays a major role in the distribution of specific syntactic elements such as clitics. More specifically, they claim that the (emphatic) discourse particle *fa* in Hausa must follow a PPh. Importantly, the distribution of this element reveals the crucial role branchingness plays in PPh formation in Hausa. To explain, the particle *fa* cannot appear between a verb and its object if the latter is not branching. Compare, for instance, the ungrammaticality of (26a) and (26b) with the grammaticality of (26c). The latter example illustrates that the following element *must* be a branching (sister – head) constituent and not just any sequence of words. More importantly, it shows that branching constituents such as *babban tebur* are mapped onto a PPh.

- (26) a. $*[\text{Ya}]_{\text{PPh}} [\text{sayi fa teburin}]_{\text{PPh}}$
 he bought table-DEF
 ‘He bought the table’

- b. *[Ya]_{PPH} [sayi fa teburin]_{PPH} [jiya]_{PPH}
 he bought table-DEF yesterday
 'He bought the table yesterday'
- c. [Ya]_{PPH} [sayi]_{PPH} fa [babban tebur]_{PPH}
 he bought big table
 'He bought a big table'

On the basis of the Hausa facts as well as additional cross-linguistic evidence, Zec and Inkelas (1990) propose the so-called *arboreal mapping*, which aims at parsing branching nodes into PPhs. The thrust of their analysis, however, is that the order of constituents in languages like Hausa requires simultaneous reference to syntactic and prosodic structure. This led them to propose a bidirectional model of interface according to which both components are derivationally related in the sense that they exist side-by-side and, as such, they can mutually affect each other (Zec and Inkelas, 1990: 378). In this respect, therefore, this approach contrasts sharply with the 'syntax-first' motto advocated by the chief proponents of the RB and EB models. Interestingly, one particular aspect of their proposal, namely, the derivational mode of interaction between the two components, has been exploited in subsequent research under the influence of new theoretical advancements that pertain to the architecture of grammar as well as to the theory of syntax per se.

To sum up, early theories of the syntax-phonology interface focused on issues such as: (i) the type of syntactic information that is visible to phonology, (ii) whether phonological rules are entitled to draw on and directly handle this information from syntax or whether there is a buffer, that is, a mapping mechanism that assists the communication by constructing prosodic units such as the PW and the PPh within which phonological rules operate, and (iii) the mode in which the two components engage in their interface, that is, serial, with syntax feeding into phonology or parallel, where both components are mutually interactive. Although these issues lie at the heart of the interface, contemporary research has shown an increasing interest not only in the mapping of the syntactic objects onto phonological ones and the mechanics of it, but also in the influence phonology exercises on the linearization of the syntactic structure, that is, the way the hierarchical structure of syntax translates into a linear order, the fixing of the order of the constituents in a syntactic string and the decision upon which syntactic elements will be pronounced or silenced. Aspects of this area of research are explored in the remainder of this chapter.

3. The Minimalist Program and the Syntax-Phonology Interface

Most approaches to the syntax-phonology interface presented above have been shaped in relation to the *Government and Binding* framework (Chomsky, 1981

and following) of the *Principles and Parameters Theory* (Chomsky and Lasnik, 1993). In the last two decades, the *Minimalist Program* (MP) (Chomsky, 1993, 1995 and following) makes two very significant assumptions with important consequences on the way the syntax-phonology interface has been viewed so far.

First, it maximizes the significance of interfaces by adopting the so-called strong minimalist thesis, which maintains that language is an optimal solution to interface conditions that the Faculty of Language must satisfy. Syntax is simply a derivational procedure that mediates between the two interfaces (i.e. the Sensory-Motor (PF) and the Conceptual-Intentional (LF)), in the sense that it constructs objects (i.e. structures) which the two interfaces process and evaluate by means of their own conditions and procedures. In such a model, syntax always precedes phonology, which now aims to interpret or, rather, process the constructions created by syntax by means of its own devices. Obviously, the significance of the syntax-phonology interface as the locus where syntactic structures are prepared so as to be processed by the Sensory-Motor component is upgraded. This preparation involves (i) the linearization of the syntactic structure and (ii) the mapping of the syntactic objects onto phonological objects. Linearization concerns the translation of the hierarchical (syntactic) structure into the relevant linear order and the fixing of the order of the constituents in a syntactic string.⁹ On the other hand, the mapping procedure deals with the organization of the elements of the string into prosodic constituents such as the ones discussed in the previous sections of this chapter.

Second, if syntax is considered to be a strictly derivational mechanism which creates constructions which it feeds the interfaces, then it is expected that effects of this derivational procedure will be present in the representations that the syntax-phonology interface constructs. In other words, theoretically speaking, it would be desirable if some properties of the syntax-phonology interface regarding the mapping procedure described in the previous sections are derived by the general architecture of the grammar and the independent principles of derivation. In the early versions of the MP (Chomsky, 1993, 1995), this was not feasible because syntax and phonology were assumed to interact only once, that is, after the application of Spell-Out, which, significantly, was assumed to apply once and after the whole syntactic derivation was completed. However, recent versions of the MP adopt the so-called *Multiple Spell-Out Hypothesis* (MSOH) (Uriagereka, 1999; Chomsky, 2000 and following), according to which Spell-Out applies not just once but iteratively, sending pieces of syntactic constructions to the interfaces for processing. This derivational approach to the interface is depicted in Figure 9.4.

The MSOH splits derivation into certain derivational domains upon the completion of which the operation of Spell-Out applies and sends its output to the interfaces for independent processing. No matter how these domains are defined (*phases*, Chomsky, 2000 and following; *c-command domains/cascades*, Uriagereka,

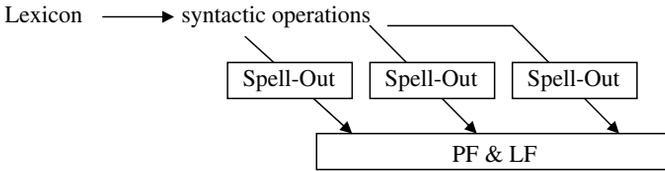


Figure 9.4 Multiple Spell-Out Hypothesis

1999; *prolific domains*, Grohmann, 2003; etc.), the MSOH presents a dynamic non-linear approach to the syntax-phonology interface because the prosodification of a given syntactic string is derived on the basis of the products of each application of Spell-Out. That is to say, each product of Spell-Out is directly mapped onto a certain constituent of the PH. There is a growing body of literature on this issue, which mainly aims to relate *phases* with higher order prosodic structure (see Elordieta, 2007 for an overview). For example, it has been proposed that sentential stress and intonational phrasing of a clause can be independently derived by mapping phases onto specific prosodic constituents such as phonological or intonational phrases (Ishihara, 2003, 2007; Kahnemuyipour, 2004, 2005; Adger, 2007; Kratzer and Selkirk, 2007, among many others).

There are some very interesting and significant issues raised by such an approach, such as (i) What is the status of EB algorithms in such a model?, (ii) How can purely phonological conditions and operations on prosodification such as binarity and, in general, rephrasing be accounted for or better accommodated within this model?, and finally (iii) Do the different products of Spell-Out communicate among each other at the syntax-phonology interface and, if yes, under what conditions? All the above questions constitute a very interesting and promising line of further research, the results of which are much awaited (see Revithiadou and Spyropoulos, 2003, 2005, 2009, for some recent findings).

Let us now return to the role the syntax-phonology interface plays in the linearization of the syntactic derivation. This is a novel, yet very significant field of inquiry about the syntax-phonology interface, which emerged from the way the MP approaches displacement phenomena. Displacement is one of the core properties of language. Syntactic elements (i.e. words, phrases) may be pronounced in a different position from the one in which they acquire their semantic (theta-) role in the clause. Displacement is considered to be a *par excellence* syntactic phenomenon and it is accounted for by means of the syntactic operation of movement. Movement as an operation relates two different positions in the syntactic configuration, namely a position in which the displaced element acquires its theta role and a position in which it licenses certain

grammatical properties. The significant development of the MP is that it adopts the so-called *Copy Theory of Movement* (see Corver and Nunes, 2007 for an overview). Movement is viewed as a copying mechanism whereby the displaced element does not move leaving behind traces, but rather it is copied and its copies occupy the relevant positions in the syntactic configuration. The motivation behind this development was the need to account for complex reconstruction data at LF (see the discussion in Chomsky 1993, 1995 and in many works thereafter), that is, cases where the displaced element causes interpretation effects not from its final position but from the intermediate positions or even the original position of movement. At first, it was assumed that the higher copy is always the one pronounced. However, empirical facts from a range of languages indicated that other copies may – under certain circumstances – also be pronounced (see the overview in Nunes, 2004).

Initially, the discussion concentrated on cases where it is the lower copy that is pronounced instead of the higher one. Franks (2000) and Bošković (1995, 2001) first identified cases where the higher copy of a movement is not pronounced because its pronunciation would violate certain conditions on the phonological make-up of the sentence (see also Nunes, 2004 for discussion and Revithiadou, 2006 for more data from second position clitics in Greek dialects). Bobaljik (2002) goes a step further and proposes an approach to movement according to which it is a strictly syntactic process which creates a series of copies at the relevant positions. This sequence of copies is nothing more than the syntactic output of movement after Spell-Out takes place. Which copy will be pronounced by PF and/or interpreted by LF is not a matter of syntax but rather of the relevant interfaces.

The approaches mentioned above explicitly state that principles of the syntax-phonology interface are responsible for the so-called *linearization of chains*, that is, the decision about which copy in a given movement output will be pronounced or not. Given this, phonology seems to have a pervasive effect on fixing the constituent order of a construction, a possibility also explored by Pesetsky (1996, 1998) within the framework of OT. These effects concern the blocking of the pronunciation of a copy because its pronunciation would violate certain prosodic conditions and/or the promotion of a certain copy because its pronunciation is the most optimal choice for the prosodic make-up of the construction. Notice that such approaches by no means imply that phonology affects syntax. They rather underline the significance of the interpretative role of phonology. Syntax as a generative device simply creates structures and phonology processes them on the basis of its own principles and constraints.¹⁰ Clearly, this line of research maximizes the autonomy of the two modules and the role of the interface, a welcome result for the architecture of the grammar.

Regarding the mechanism of copy linearization in the syntax-phonology interface, we identify two possible versions. The weak version states that in a

movement operation the default case is to pronounce the higher copy. Pronunciation of other copies is allowed only if pronunciation of the higher copy violates certain well-formedness conditions on the prosodic make-up of the construction. On the other hand, the strong version states that the choice upon which copy will be pronounced depends solely on which copy is more optimal – from a phonological point of view – to be pronounced in a given construction. The difference between these two theses is that the weak version promotes a certain copy, the higher, perhaps as a residue of the syntactic derivation (see Nunes, 2004, for some proposals), whereas the strong version leaves the interface free to operate on its own principles. Interestingly, the bulk of work done on copy linearization silently or explicitly adheres to the weak version.¹¹ Research on this topic is rather poor and subsequent work should focus on issues such as (i) what promotes the higher copy at the interface and (ii) what type of conditions determine the pronunciation of copies either by blocking or by promoting them. In the following section, we review the results from a study from Greek, which show how both the blocking of the pronunciation of a copy and the promotion of a certain copy for optimality reasons determine copy realization and subsequently word order.

4. Case Study: The Linearization of Subjects in Greek Clauses

4.1 Setting the Stage: Distribution and Derivation of Greek Subjects

The syntax of Greek subjects has been one of the most studied topics of Greek syntax. The major challenge is that their distribution in the clause is quite complex even in discourse neutral situations with default intonation in all new information contexts. Thus, subjects in Greek (i) may be either pre-verbal or post-verbal in transitive constructions (SVO ~ VSO, with SVO being the preferred option by far), (ii) are preferred to be post-verbal in intransitive constructions and (iii) are obligatorily post-verbal in subjunctive clauses (see Philippaki-Warbuton, 1985 and Laskaratou, 1998, among many others). Most of the studies assume that each case involves a different derivation with pre-verbal subjects being left-dislocated elements and post-verbal ones occupying the relevant theta position inside the *v*P (cf. Philippaki-Warbuton, 1987; Alexiadou and Anagnostopoulou, 1998; Spyropoulos and Philippaki-Warbuton, 2001).

In a recent study, Spyropoulos and Revithiadou (2009) take into consideration the prosodic properties of Greek subjects and propose an interface analysis which accounts for their distribution by means of a single syntactic derivation which creates a sequence of subject copies and the linearization of this subject chain at the syntax-phonology interface after Spell-Out. They extensively argue against the left-dislocation analysis of pre-verbal subjects on the basis of their prosodic,

syntactic and semantic properties (see also Roussou and Tsimpli, 2006) and claim that Greek subjects always move from their theta position to the EPP Spec, TP position. The syntactic output of this movement is a sequence of copies, one in Spec,TP and another in the theta position within the *v*P. More importantly, they propose that it is the linearization of this subject chain at the syntax-phonology interface which derives the distribution of subject in a clause. The syntax-phonology interface processes this sequence of copies (the syntactic output of subject movement) and promotes the pronunciation of one of the copies as the most optimal PF output on the grounds of PF conditions and constraints (adjacency conditions, prosodic well-formedness constraints, stress assignment etc.). For instance, if we assume that the default case is that the highest copy is pronounced, the linearization of the chain in favour of this copy will be captured by the constraint PRONOUNCE HIGHEST (PHIGH) under the assumption that PF conditions are formulated in OT terms (27).

- (27) PRONOUNCE HIGHEST (PHIGH) (based on Franks, 2000)
The higher copy must be pronounced.

PHIGH, as a *positional faithfulness* constraint (Beckman, 1997), promotes the pronunciation of the head of a chain and its effects are only detectable when it is high-ranking with respect to its non-positional correlate PRONOUNCE ('A copy must be pronounced'), which indiscriminately advocates the pronunciation of every copy available in the chain, unless it is blunted by another PF constraint such as PHIGH. Interestingly, this is the ranking responsible for the emergence of discourse neutral SVO constructions. Needless to say, there are several PF-constraints involved in the evaluation of the optimal copy, the interaction of which, as encoded in a language-specific constraint hierarchy, provides a powerful tool for capturing the variation in the linearization of chains. In what follows, we demonstrate how such an interface analysis, which exploits the full power of PF-constraints, provides a straightforward analysis for the distribution of Greek subjects in subjunctive and indicative clauses.

4.2 Subjunctive Clauses

Subjunctive in Greek consists of the verb form and the subjunctive particle *na*, which resides in the head of a Mood functional category (Philippaki-Warbuton, 1998, among others). Overt subjects in subjunctive clauses are never pre-verbal when rendered with neutral intonation.¹² They also never appear between the particle *na* and the verb form, since the particle is always proclitic to the verb form. In fact, overt subjects can only be post-verbal in subjunctive clauses as shown in (28).

- (28) (*o kóstas) na (*o kóstas) fái o kóstas
 the Kostas-NOM SUBJ the Kostas-NOM eat-3SG the Kostas-NOM
 to axláði
 the pear-ACC
 ‘Kostas should/may eat the pear’

This distribution can be easily accounted for by the interface approach to copy linearization presented above. Subject movement targets the EPP Spec,TP position where the copy is structurally situated in-between the particle *na* and the verb form which raises to the T head as shown in (29).

- (29) [_{MP} [_M na] [_{TP} o kóstas [_T fái] [_{VP} o kóstas fái to axláði]]]

Philippaki-Warburton and Spyropoulos (1999) argue that the particle *na* and the verb form are subject to a post-syntactic merger, which operates on string adjacent elements (see Marantz, 1988) and which is responsible for the surfacing of *na* as a proclitic to the verb. The pronunciation of the higher copy of the DP-subject *o kóstas* in Spec,TP position would be fatal for the merger. The silencing of this copy suggests that some other constraint promotes the pronunciation of the lowest copy in the chain. Spyropoulos and Revithiadou (2009) attribute this effect to a constraint STRICT ADJACENCY, stated in (30), which outranks PHIGH: STRICT ADJACENCY » PHIGH » PRONOUNCE.

- (30) STRICT ADJACENCY (SA)
 Elements liable to post-syntactic merger must be strictly adjacent.

An ungrammatical output such as *na o kóstas fái ~~o kóstas~~ . . .* will be ruled-out by SA, leaving the string with the lowest copy pronounced, *na ~~o kóstas~~ fái o kóstas . . .*, as the optimal output. PRONOUNCE will be violated in both cases since not all copies of the string have the chance to be pronounced.

4.3 Indicative Clauses

In Greek indicative transitive clauses, both SVO and VSO orders are felicitous answers to questions requiring an all new information answer as shown in (31).

- (31) Q: tí éyine? /tí néa?
 what-ACC happen-PAST.3SG/what-NOM new-PL.NOM
 ‘What happened?/What’s up?’

- A1. o kóstas éfaje to axláði
the kóstas-NOM eat-PAST.3SG the pear-ACC
- A2. éfaje o kóstas to axláði
eat-PAST.3SG the kóstas-NOM the pear-ACC
‘Kostas ate the pear’

Interestingly, SVO is the optimal and most frequent answer in such cases (Keller and Alexopoulou, 2001; Laskaritou, 1989, 1998). However, the availability of the VSO order suggests that the requirement to pronounce the highest copy is violable in Greek; otherwise, outputs with the lower copy pronounced, which derive the VSO order, would have been considered ungrammatical. The syntactic derivation of the transitive sentence in (31) is given in (32), together with its possible linearizations.

- (32) [_{CP} [_{TP} o kóstas [_T éfaje] [_{vP} o kóstas éfaje [_{vP} éfaje to axláði]]]]
a. [o kóstas] éfaje ~~o kóstas~~ to axláði preferred
b. ~~o kóstas~~ éfaje [o kóstas] to axláði less preferred

With the completion of the *vP*-phase, Spell-Out sends its complement domain, that is, the VP, to the interfaces for processing. This contains the DP-object *to axláði*, which is mapped onto a PPh. In the next phase up, the verb moves to T and the DP-subject *o kóstas* moves to the EPP Spec,TP position. Spell-Out of the CP-phase sends its complement domain, that is, the TP and the residue of the *vP*-phase to the interfaces. This material includes two copies of the DP-subject *o kóstas* and a copy of the verb. Thus, the processing of this material by the syntax-phonology interface has to linearize the syntactic output of the subject movement, which consists of the two copies of the DP-subject *o kóstas* by promoting one of the two copies for pronunciation. This is done in the following way.

Let us assume that during Spell-Out higher order prosodic structure is assigned. Although various analyses may differ in their technical details, they all converge in accepting the spell-out domain of a phase as the domain of a PPh constituent. Spyropoulos and Revithiadou (2009) employ an EB algorithm which operates on the edges of syntactic phrases to map them into PPhs by means of ALIGN-R (XP, R; PPh, R) (Selkirk, 1995a, 2000 and following; Truckenbrodt, 1995, 1999, among others). Following Kratzer and Selkirk (2007), they assume that both phonological phrasing and sentential stress are assigned at root spell-out, that is, after the whole derivation is completed. Based on this assumption, two possible outputs are predicted for the input material of the CP-phase *o kóstas éfaje o kóstas*.

- (33) *p-phrasing in CP-Phase*
 a. [o kóstas]_{PPH} éfaje o kóstas
 b. o kóstas [éfaje o kóstas]_{PPH}

The phrasing in (33a), which results in the SV order, has a shortcoming compared to the one in (33b), which promotes the VS order: all elements of the string are contained into the PPh in the latter phrasing but not in the former. We assume that in (33a) the verb is left stray and either joins the PPh of the subject in a later application of the phrasing algorithm or rephrases so as to join the PPh which contains the object (i.e. the interface product of the processing of the complement domain of the *vP*-phase which has already been Spelled-Out and processed).¹³

The non-optimal phrasal shape of (33a) is due to a violation of the archetypical constraint NoSTRAY (34), a member of the PARSE family of constraints.

- (34) NoSTRAY
 Morphosyntactic material should be prosodically parsed.

This constraint opts for a parsimonious parsing such as *o kóstas [éfaje o kóstas]_{PPH'}*, where the V is immediately prosodified after Spell-Out, over the non-parsimonious phrasing [*o kóstas*]_{PPH} *éfaje o kóstas*. Thus, the VSO order emerges when the constraint NoSTRAY outranks PHIGH, that is, the constraint that favours the pronunciation of the head of the chain as shown in (35). The opposite ranking is responsible for the emergence of the SVO order.

- (35) a. ALIGN-R, NoSTRAY » PHIGH → VSO
 b. PHIGH » ALIGN-R, NoSTRAY → SVO

Regardless of which phrasing and, consequently, word order is chosen, the phrased string will also have to be assigned sentential stress. In Greek, sentence stress prominence is on the rightmost element of the PPh (RIGHTMOST-PPH) and the IP (RIGHTMOST-IP). For both the SVO/VSO orders, this element is the DP-object. Interestingly, the very same set of PF-constraints yields somewhat different results for agent-less intransitive clauses. More specifically, in intransitive constructions, the VS order seems to be preferred. The interface approach to subject chain linearization can also account for this difference.

The derivation of a typical intransitive construction such as the one in (36) is presented in (37).

- (36) írthe o kóstas
 come-PAST.3SG the Kostas-SG.NOM
 'Kostas came'

(37) [_{CP} [_{TP} o kóstas [_T írθe] [_{vP} írθe [_{VP} írθe o kóstas]]]]

The movement of the DP *o kóstas* to the Spec,TP position creates a chain with a pre-verbal copy at the Spec,TP and a post-verbal copy inside the VP. We assume that agent-less intransitive predicates involve a Spec-less *vP*, which is also subject to Spell-Out, despite its weak phase status (see Legate, 2003 and Kratzer and Selkirk, 2007). Thus, the example in (37) involves a two-phase derivation. Spell-Out of the weak *vP*-phase sends its complement domain, that is, the VP, which contains the lower copy of the DP *o kóstas*, to the syntax-phonology interface for prosodification, where the DP will form a PPh: [*o kóstas*]_{PPh}. Accordingly, the material of the CP-phase *o kóstas írθe* will be Spelled-Out and prosodified as shown in (38): [*o kóstas*]_{PPh} írθe.

(38) [_{CP} [_{TP} o kóstas [_T írθe] [_{vP} írθe [_{VP} írθe o kóstas]]]]

- a. weak *vP*-phase Spell-Out: [_{VP} írθe o kóstas] → [*o kóstas*]_{PPh}
 b. CP-phase Spell-Out: [_{TP} o kóstas [_T írθe] [_{vP} írθe] → [*o kóstas*]_{PPh} írθe

In this prosodification, the verb remains trapped in-between two PPhs and it will have to be prosodically incorporated during rephrasing or, in general, a later application of the phrasing algorithm. Interestingly, regardless of which copy will be phonetically realized, the end-result in both cases is a phrasing in which the verb has to adjoin to the PPh of the DP-subject or phrase by itself.

- (39) a. [*o kóstas*]_{PPh} írθe *higher copy is pronounced*
 b. írθe [*o kóstas*]_{PPh} *lower copy is pronounced*

An important question that arises at this point is what hinders the effect of PHIGH so that (39b) is grammatical (and more optimal) in Greek. Spyropoulos and Revithiadou (2009) consider the possibility that the final shape of the string is determined by the constraints that are responsible for the rightmost sentence stress rule, that is, RIGHTMOST-PPh/IP. In contrast to transitive constructions, in which sentence stress falls on the DP-object, in intransitive constructions sentence prominence varies depending on which copy of the DP-subject is pronounced: in VS orders stress is on the DP-subject whereas in SV orders stress is on V. This entails that in (39a) sentence stress can only be assigned when all material is prosodically organized. When this is done, the V will receive sentential stress. If, however, the lower copy is selected as in (39b), sentential stress can be assigned right away. The structures in (40) illustrate the position of sentence stress in both word orders.

(40) *sentence stress in SV/VS orders*

$$\begin{array}{l} \{ \quad \quad x \} \quad \{ \quad \quad x \} \quad \text{IP} \\ [[\quad] x] \quad [\quad [x]] \quad \text{PPh} \\ (x) (x) \quad (x) \quad (x) \quad \text{PW} \end{array}$$

a. o kóstas írthe b. írthe o kóstas

To conclude, we have shown that several PF-constraints, appropriately ranked, for example, SA » {ALIGN-R, NoSTRAY, RIGHTMOST-PPh/IP} » PHIGH » PRONOUNCE, determine phonological phrasing in Greek and, at the same time, play a role on which copy of the subject will be pronounced. This result has more than abstract interest because, in essence, it derives from an analysis that postulates one constraint ranking to account for both phonological phrasing *and* word order. Only a copy theory of movement can provide the necessary leeway to the PF interface to dynamically process the syntactic output, thus achieving considerable descriptive and explanatory economy compared to other theories.

5. Conclusions

This chapter explores basic aspects of the syntax-phonology interface. We have shown that research in this area has led to important results. First, the interface is mediated by units of prosodic structure which cross-linguistic research has established as serving as domains for the application of phonological rules. Second, although it is not agreed on how exactly the mapping of syntactic and prosodic structure is pursued, it is fairly certain that the mapping is not always isomorphic. It is still an open question, however, whether syntax impinges on phonology or the interface is bidirectional. In search of an answer to this question, we explored some recent advances on the architecture of grammar, which yield some interesting results for the interface. In particular, we focused on the issue of linearization and explored the effect of independently motivated phonological constraints on determining not only the prosodic structure but also the word order of a given syntactic string.

Our case study was the distribution of Greek subjects. In the realm of the copy theory of movement, we proposed that the movement operation creates a sequence of copies, which constitutes the syntactic output of this operation, and PF fixes their surface distribution on the basis of independently existing phrasing constraints. In fact, PF invokes a constraint system to account for both phonological phrasing and word order. It is possible that this approach can tackle some, if not all, of the data put forward in favor of the bidirectional model of the interface such as the clitic *fa* in Hausa.

It is far from certain whether the questions posed in the past 40 or so years of research in this area have been answered in a satisfactory way. It is certain that many more wait to be answered or even addressed. The discussion in this chapter has – hopefully – made clear that progress in this area can be achieved only when independent developments in syntax and/or phonology go hand-in-hand with a holistic and unified view on the architecture of the Grammar.

6. Acknowledgements

We wish to thank the editors and two anonymous reviewers for their insightful comments. All errors are of course our own.

7. Notes

1. In general, boundaries encode higher order prosodic structure on a string of segments and, as such, delimit the domain of application of phonological rules (McCawley, 1968; Selkirk, 1972, 1980b, 1981a). Essentially, boundaries signify an indirect mode of interaction between the components of grammar: phonological rules can either refer to boundaries in their structural description or be blocked by them, but they can never refer directly to syntactic edges.
2. This debate was explored in a special volume of *Phonology* edited by Kaisse and Zwicky (1987).
3. Odden (1990) allocates syntax-sensitive rules to the lexical component thus reserving the post-lexical component for rules which ‘are blind to nonphonological structures’ (Odden, 1990: 268). This assumption has a rather important consequence: Syntactic structure is available to phonology both lexically and post-lexically, hence in the architecture of grammar the lexicon and the syntax run in parallel since the former component may well contain syntax-sensitive rules.
4. See Hockett (1955) and Haugen (1956) for the notion of hierarchical structure in pre-generative literature.
5. The CG is the most debatable constituent of the PH. It was originally introduced by Hayes (1989b [1984]) and later adopted and further established by Nespor and Vogel (1986) and subsequent work (Nespor, 1999; Vogel, 1991, 1997, 1999, 2009; Kabak and Vogel, 2001). A number of studies argue against its inclusion in the PH on the basis of typological evidence regarding the prosodic organization of clitics (Zec, 1988, 1993; Booij, 1988, 1995, 1996; Inkelas, 1989; Zec and Inkelas, 1991; Selkirk, 1995a; Peperkamp, 1997, among others).
6. The two theories differ in the mapping algorithm employed for the construction of PWs. Nespor and Vogel (1986) propose a different mapping rule for the construction of PWs, whereas Selkirk (1986) introduces the EB mapping for the construction of both PWs and PPhs. In this sense, the latter theory appears to be more unified than the former. Due to space limitations, we cannot provide a comprehensive account of each model here. The two models are examined in detail and compared both at the empirical and theoretical level in Revithiadou (to appear).
7. Vogel altered her position regarding the status of RS as a PPh-domain rule. Instead, she claims that RS in Tuscan Italian ‘seems to apply throughout sentences, without

regard to their syntactic (and phonological) constituency' (Vogel, 1997: 66). For more problems regarding RS as a diagnostic for PPhs see also Monachesi (1999).

8. Branchingness has been shown to play a role in the phrasal phonology of a wide variety of languages such as Hausa (Zec and Inkelas, 1990), Korean (Cho, 1990; Jun, 2003), Mende (Cowper and Rice, 1987), Mandarin (Cheng, 1987), Kinyambo (Bickmore, 1990), various Romance languages (Ghini, 1993; Prieto, 1997; Elordieta et al. 2003; Sandalo and Truckenbrodt, 2001) and Tiberian Hebrew (Dresher, 1994).
9. The procedure of linearization was extensively discussed and investigated within the *Antisymmetry of Syntax* program initiated by Kayne's (1994) seminal work. Kayne proposed that the linearization of the hierarchical syntactic string is derived by the *Linear Correspondence Axiom*. The exact status and place in the architecture of grammar of this Axiom is an open issue, although after Chomsky (1995) it is generally assumed that it is a property of the syntax-phonology interface.
10. For earlier versions of 'filtering' effects of phonology on syntax see Ross (1967) and Pullum and Zwicky (1988).
11. See Franks (2000) and Bošković (2001) on Serbocroatian clitics, Rice and Svenonius (1998) for V2 effects in Norwegian, and Revithiadou (2006) on second position clitics in certain Greek dialects. But, see Revithiadou and Spyropoulos (2006) and Revithiadou (2008) for an approach to the typology and diachrony of Greek cliticization which evaluates the pronunciation of the relevant clitic copies solely on the grounds of independently formulated well-formedness prosodic templates. See also Richards (2006) and Trinh (2009) for some other case studies.
12. Overt subjects may appear before the particle *na* but when they do so, they must be rendered with a contrastive focus (i) or a topic (ii) intonation (indicated with ||):

- | | | | | | | |
|---------|-----|---|------|----------|-----|----------|
| (i) a. | o | kóstas | na | érθi | | |
| | the | Kostas-NOM | SUBJ | come-3SG | | |
| | | 'KOSTAS should come' | | | | |
| b. | o | kóstas | na | fái | to | axláði |
| | the | Kostas-NOM | SUBJ | eat-3SG | the | pear-ACC |
| | | 'KOSTAS should eat the pear' | | | | |
| (ii) a. | o | kóstas | na | érθi | | |
| | the | Kostas-NOM | SUBJ | come-3SG | | |
| | | 'As for Kostas, he should come' | | | | |
| b. | o | kóstas | na | fái | to | axláði |
| | the | Kostas-NOM | SUBJ | eat-3SG | the | pear-ACC |
| | | 'As for Kostas, he should eat the pear' | | | | |

In such cases, they constitute (contrastive) foci or topics and they involve a different derivation, namely movement to a Focus projection or left dislocation, respectively.

13. Alternatively, the verb could also be phrased by itself under the pressure of EXHAUSTIVITY, as aptly suggested by an anonymous reviewer. In this case, the PPh of the verb would possibly become subject to restructuring at root Spell-Out when the remainder of the structure becomes available for prosodification. At this point, it remains an open question whether such elements are left to be prosodified in later stages of the derivation or are phrased and further subjected to rephrasing at the end of the derivational cycle. We opt for the former scenario here merely on grounds of simplicity: stray elements entail less structure and hence are more available to the parsing conditions in later stages of the derivation.

10 Sign Language Phonology

Chloë R. Marshall

Chapter Overview

1. Introduction	254
2. Processing Language in the Visuo-Gestural Modality	255
3. The Phonology of Sign Languages	259
4. Case Study: Segmenting the Sign Stream	272
5. Conclusion: Current Controversies in Sign Language Phonology	274
6. Acknowledgements	276
7. Notes	276

1. Introduction

The vast majority of the world's languages are produced by the mouth and perceived by the ears. However, certain languages make use of a different modality: they are produced by the hands and perceived by the eyes. Despite their very different method of transmission and reception, such languages, known as sign languages, are natural languages just as spoken languages are. If they were merely pantomime and gesture, then sign languages would arguably be mutually intelligible, and very easy for non-native users to learn, but that is not the case. Nor is there just one universal sign language; instead, there are many distinct sign languages in the world. Sign languages originate either where deaf people come into contact and need to communicate, or where an existing sign language is introduced to a deaf community from outside and then diverges from its source. Importantly, sign languages are not based on spoken languages, even though they have contact with them.

Only in the last 50 years have we understood that the linguistic complexity of sign languages rivals that of spoken languages. Much of the credit for this breakthrough goes to William Stokoe, who carried out pioneering research into

the structure of American Sign Language (ASL) in the late 1950s and 1960s. Stokoe's ground-breaking discovery was that just as spoken languages have duality of patterning, whereby meaningless units are organized into meaningful ones, so too do sign languages. His research, subsequently confirmed by many others, showed that sign languages have all the hallmarks of fully developed, natural and independent languages, i.e. syntax, morphology, semantics and phonology.

As we shall see in this chapter, the definition of phonology as the study of 'the organisation of the sounds of speech' (Roca and Johnson, 1999: 3) is too narrow. A more accurate definition is that 'phonology is the level of linguistic structure that organises the medium through which language is transmitted' (Sandler and Lillo-Martin, 2006: 114). Furthermore, studying sign languages makes us question whether theoretical constructs such as the feature, segment and syllable, which are so widely accepted in the phonological literature, are truly valid universally.

My aim in this chapter is to convey something of the excitement of studying sign language phonology, and to demonstrate what sign languages can contribute to our understanding of phonology. We will revisit some of the topics of previous chapters, but this time from the point of view of sign languages. Some aspects of sign language phonology will seem surprisingly familiar after reading those chapters, but others will appear very different. Finally, we will apply this background knowledge to a particular phenomenon: how signers segment the sign stream in order to recognize individual signs. The best-studied sign language is ASL, and so this language will predominate in what follows, but for practical reasons the photographs depict signs from British Sign Language (BSL)¹, the language used by the Deaf² community in the United Kingdom. Few other sign languages have yet been studied in detail, and therefore cross-linguistic comparisons are difficult, but I will bring in examples from languages such as Sign Language of the Netherlands and Al-Sayyid Bedouin Sign Language. However, let us first consider what languages in the visuo-gestural modality look like and how they are processed.

2. Processing Language in the Visuo-Gestural Modality

2.1 Simultaneity Versus Sequentiality

The hands obviously play a central role in producing sign language. Hands move in a signing space that is roughly limited to the space from the hips to just above the head. Hands and arms are bigger than the lips and tongue, and therefore move more slowly: fluent signers produce only 2–3 signs per second,

compared to the 4–5 words per second produced in running speech. Despite this difference, the proposition rate for sign and speech is the same, at roughly one proposition every 1–2 seconds (Bellugi and Fischer, 1972).

Hands, however, are not the only articulators in sign languages. Non-manual articulators, namely the different parts of the face (particularly the eyebrows, eyes and mouth) and the torso, are also important. These articulators can be controlled independently of one another, making it possible for signers to produce separate pieces of linguistic information simultaneously. This does not pose a problem for perception because the eyes have excellent spatial resolution, and can process the different pieces of linguistic information when they are produced simultaneously on different parts of the body. In contrast, the eyes have poor temporal resolution: whereas the human ear can distinguish two stimuli presented only 2 milliseconds apart, two visual stimuli need to be presented a minimum of 25–30 milliseconds apart if they are to be distinguished (Brentari, 2002).

These differences in production and perception between the two modalities mean that signed and spoken languages use sequential (linear) and simultaneous (non-linear) processing to different extents, and herein lies the solution to the paradox of an equivalent proposition rate in the two modalities. Although some information in spoken language is processed non-linearly (e.g. formant transitions in the early part of a vowel can signal both what the vowel is and what the preceding consonant was (Stevens, 1998)), spoken words unfold over time in a linear manner. Sign languages, in contrast, present a much greater proportion of their phonological material in a non-linear fashion (although there is a linear aspect to signs in that, for example, the initial and final locations of a sign are often different). This is the case for morphosyntactic information too. In spoken languages, bound morphemes tend to be added to their bases as prefixes, suffixes or infixes, although there are exceptions to this, such as umlaut and tonal morphemes. Sign languages, however, have a strong preference for morphemes to be presented non-linearly. An example is aspect in ASL, which is conveyed by the movement of the verb: different rates, rhythms, degrees of tenseness and pauses modify the meanings of basic signs such as LOOK-AT and BE-SICK³, and these movements are integral to the sign (Klima and Bellugi, 1979). This morphological property has led researchers to talk about spoken languages being processed ‘horizontally’ and sign languages being processed ‘vertically’ (Brentari, 2002).

2.2 Iconicity Versus Arbitrariness

A further difference between signed and spoken languages is the different degree to which they make use of iconicity. The term ‘iconicity’ refers to a direct or

transparent relationship between form and meaning. According to de Saussure (1916[1959]), arbitrariness between form and meaning is a fundamental property of language. For example, the fact that tree is 'tree' in English, 'arbre' in French and 'Baum' in German bears no relation to the meaning of tree – the different phonological forms are purely arbitrary. Fairly obviously, it is easier for visual languages to represent the visual forms of objects, and therefore to be iconic, than it is for spoken languages to represent the auditory forms of objects (few objects make characteristic noises, and even fewer make noises that can be easily mimicked by the human vocal apparatus). Sign languages therefore use iconicity to a greater degree than spoken languages, but not all signs are iconic. Even for those that are, there is a frequent arbitrariness about which aspects of the referent they choose to encode iconically. Take, for example, the sign COFFEE in two sign languages, BSL and ASL. BSL encodes the drinking aspect of the meaning of coffee, whereas ASL encodes the grinding of the coffee beans (Sutton-Spence and Woll, 1999). It is a similar story for the onomatopoeic words used to encode sounds in spoken languages. Words like 'miaow' and 'nee nar' represent for English speakers the sounds that cats and sirens make, but Japanese cats go 'nyaanyaa', and Spanish sirens go 'tut tut'. Thus, even iconicity needs to be conventionalized, whatever the modality.

One might expect the transparent mapping between form and meaning in iconic signs to aid Deaf people's processing of sign language, and a recent study has indeed shown this to be the case for lexical retrieval (Thompson et al., 2009). In an elegant experiment, adult ASL users were shown a picture followed by a sign, and the picture was one of two types: it either made the iconic property of the sign highly visually salient, or it did not make that property salient. For example, there were two pictures of a hearing aid. The sign for HEARING AID in ASL is produced by the curved index finger and thumb tapping just behind the ear, to represent the hearing aid being held in position there. In the high salience condition, the picture was a close-up of a hearing aid fixed in just that position. In the low salience condition the picture was of a smaller aid sitting inside the ear. When the pictures were highly salient, signers were faster to respond that the sign matched the picture, indicating that iconicity was helping signers to access the sign in their lexicon. Thompson and her colleagues conclude that 'true arbitrariness [of language] may be the result of language modality, rather than a stand-alone feature of language'.

2.3 Infants' Modality-Independent Bias for Language

It has long been known that in their first year of life, before they are able to produce any words, hearing infants pay attention to linguistic sounds in their environment and become experts at discriminating and categorizing those

sounds. Studies have shown that 4 month olds prefer speech to silence and white noise, and that 9 month olds prefer listening to singing rather than instrumental music (see Jusczyk, 1997, for a review). Nor is this speech bias just a bias towards complex sounds: 2–7 month olds prefer listening to spoken words over complex sine-wave analogues⁴ of those same words (Vouloumanos and Werker, 2004).

This bias towards speech actually appears to be a bias towards language in general, rather than towards acoustic properties. Krentz and Corina (2008) showed hearing 6-month-old babies, who had never seen any sign language before, a Deaf native signer using ASL to describe various everyday scenarios and the same woman performing everyday pantomime sequences (e.g. putting on make-up, frying an egg). When shown ASL and pantomime side-by-side on two TV monitors, babies looked for longer at the woman using ASL. In a follow-up experiment, Krentz and Corina then recorded the same woman performing pantomime and using ASL, but this time using the point-light technique. In this technique, lights are attached to the forehead, chest and major parts of the torso, and the recording is carried out in a dark room – the video camera catches only the movement and location of the lights, and no information from the hands or face. This time the 6 month olds preferred to look at the pantomime. The authors interpreted their results as follows: in sign languages, the hands and face contain a lot of information that is highly salient to infants, much more so than pantomime. Infants attend to this salient information. When that information is missing, the larger movements of pantomime (recall that movements in sign languages are confined to a well-defined space) capture children's attention. The nice twist in the design of this study is that the authors have shown that the language bias does not emerge from experiencing language in a particular modality: hearing babies prefer sign language over pantomime even when they have never seen sign language before.

2.4 Categorical Perception

Just as there is categorical perception (CP) for certain classes of speech sounds, so there is for certain aspects of sign language. Emmorey et al. (2003) tested CP in native signers of ASL using discrimination and categorization paradigms with computer-generated images of signs. Signers demonstrated CP for handshape but not for location. CP for handshape is likely to be a robust finding, given that it has been replicated with a wider range of handshapes and with natural rather than synthetic stimuli (Baker et al., 2005).

This work has been extended to infants. Using one of the handshape continua from their adult study, Baker et al. (2006) found that 4-month-old *hearing* infants were able to discriminate between handshapes on the basis of linguistic

category membership, while 14 month olds failed to do so. The different performance of the two age groups is significant here, because it is identical to the initial capacity and classic developmental shift in infant categorical discrimination of native and non-native phonemes in speech. It appears then that the young human brain is specialized to detect the underlying contrasting patterns in language rather than simply processing the auditory perceptual features of sounds.

In this section, we have seen that with regard to infants' facility for tuning in to language, the similarities between signed and spoken languages are striking. We have also seen two ways in which modality affects language processing, and makes the processing of sign language appear rather different to that of spoken language – namely, its greater reliance on simultaneous, or vertical, processing, and its greater use of iconicity. A consideration of modality will continue to be important as we explore the phonology of sign languages in the next section. Indeed, the issue of just how far into the phonology the effects of modality reach is one that is hotly debated (see Sandler and Lillo-Martin, 2006, as well as the papers in Meier et al., 2002).

3. The Phonology of Sign Languages

3.1 Phonological Parameters

The basic analysis of sign structure dates back to the work of Stokoe (1960), who demonstrated that signs are not unanalysable gestures, but are instead made up of three basic phonological categories or 'parameters' that he likened to phonemes: handshape, movement and location. These terms are fairly self-explanatory. 'Handshape' denotes the particular shape that a hand makes in a sign, and handshapes vary in the number of fingers that are selected and how those fingers are flexed or extended. There are two classes of movement: 'path' movements, which involve movement of the hand and arm, and 'hand-internal' (sometimes termed 'local') movements, that involve just the fingers or wrist. Signs have either one or both of these types of movement. Signs can be produced in a neutral location in front of the signer, or on the non-dominant hand, face or torso.

The primary source of evidence for the claim that handshape, movement and location are parameters is the existence of minimal pairs for each. Just as pairs such as *tap* and *cap* show that /t/ and /k/ are phonemically contrastive in English, so pairs of signs exist which differ just in one parameter, demonstrating the contrastiveness of that parameter. For example, one of the BSL signs for DOCTOR and the standard sign for MORNING have the same location and movement, but differ in handshape (see Figures 10.1a and 10.1b).



Figure 10.1a BSL DOCTOR



Figure 10.1b BSL MORNING



Figure 10.1c BSL ARMY

It is worth clarifying at this point that these parameters are meaningless, and therefore true phonological units. This is not to say that certain movements, locations and handshapes *cannot* bear meaning – indeed they can. The ‘A’ handshape (a fist with the thumb outstretched) is used in many BSL signs that have a positive meaning, such as GOOD, FANTASTIC, CORRECT, FAVOURITE, CONGRATULATIONS and PROUD. Similarly, the head is the location for the semantically related BSL signs THINK, KNOW, NOT-KNOW, IDEA, UNDERSTAND, CLEVER, REMEMBER and FORGET (Brien, 1992).⁵

Nevertheless, just because signs *can* be analysed into different parameters does not necessarily mean that these parameters are psychologically real: what evidence is there that signers’ brains do actually analyse signs in this way? One piece of evidence comes from ‘slips of the hand’, which are the equivalent of ‘slips of the tongue’ in speech. Slips of the tongue, such as ‘teep a kape’ (instead of ‘keep a tape’) and ‘available for exploitation’ (instead of ‘available . . .’) provide evidence that phonemes are discrete units that may be mis-selected during speech production (Fromkin, 1971). Analogous selection errors occur in sign production: signers have been reported to make errors on all three major parameters (Newkirk et al., 1980). Furthermore, as in speech, these errors are constrained in that they almost always result in a phonologically possible sign.

This provides evidence that handshape, location and movement are discrete phonological units.

Since Stokoe's work, it has been proposed that there are two additional phonological parameters: 'orientation' (Battison, 1978) and 'non-manual features' (Liddell, 1980). Orientation refers to how the palm of the hand is oriented relative to the body. Minimal pairs exist which contrast solely in orientation, for example, MORNING and ARMY in BSL (see Figures 10.1b and 10.1c). Non-manual features refer to the linguistic use of body parts other than the hands, and include facial expression, eye gaze and the position and movement of the head and upper body. These have been compared to intonation in spoken languages (Crasborn, 2006). At the lexical level, some signs contain an oral component that is arguably integral to the sign's meaning and occurs at the same time as the hand movement. For example, in the BSL sign REALLY, where the active hand is brought down onto the passive hand and contact is made as the mouth closes (Figures 10.2a and 10.2b). Another example is the BSL sign MIND-BLANK, where hands move apart and close as the stuck-out tongue is drawn back into the mouth with an intake of breath. The term 'echo phonology' has been used to describe the subset of mouth gestures that parallel and appear to be driven by manual movements (Woll, 2001).



Figure 10.2a BSL REALLY (start of sign)



Figure 10.2b BSL REALLY (end of sign)

3.2 Constraints on Sign Language Structure

Just as spoken languages place constraints on how phonological units can be combined, so do sign languages place constraints on particular parameters. The most important of these constraints are the following:⁶

- **Place Constraint** (Battison, 1978): There can be only one major body area (i.e. head, trunk, arm or hand) specified in a sign. A change of location is permissible within a body area, for example, the change of location in ASL DEAF from the ear to the chin.
- **Dominance Constraint** (Battison, 1978): If the hands of two-handed sign do not share the same handshape, then one hand must be passive while the active hand moves, and the passive hand can take only one of a small set of handshapes.
- **Symmetry Constraint** (Battison, 1978): If both hands move independently, then they must be specified for the same location, the same handshape, the same movement (whether this be performed simultaneously or asynchronously), and the specification for orientation must be either symmetrical or identical.

- **Selected Fingers Constraint** (Mandel, 1981): Only one group of fingers may be selected in a sign, selected fingers being those that are ‘active’ in that they move during the production of a sign or touch the body.

Signs which violate these constraints tend to be morphologically complex. For example, the BSL sign for BELIEVE is a compound of the signs THINK^TRUE, and breaks the place constraint by being located first at the forehead and then on the palm of the non-dominant hand (Brien, 1992). The BSL sign PARENTS is a compound of the signs MOTHER^FATHER, and breaks the selected fingers constraint by using first the index, middle and ring fingers, and then just the index and middle fingers (Brien, 1992).

3.3 Markedness in Sign Language Phonology

The notion of ‘markedness’, familiar from spoken language phonology, also has a counterpart in sign language phonology. Certain handshapes are unmarked in that, in true Jakobsonian fashion (Jakobson, 1968), they are frequent cross-linguistically, are easy to articulate, are acquired first by children and are resistant to loss in aphasia (Sandler and Lillo-Martin, 2006). Furthermore, only unmarked handshapes can be taken up by the passive non-dominant hand, as required by the Dominance Constraint. The exact set of unmarked handshapes may vary very slightly from language to language, and this issue requires cross-linguistic research, particularly for non-western sign languages. For BSL, Sutton-Spence and Woll (1999) propose that the set of unmarked handshapes consists of ‘B’, ‘5’, ‘A’ and ‘G’.⁷ These are illustrated in Figure 10.3a. When these are compared to some of the marked handshapes, as shown in Figure 10.3b, the former are maximally distinct, and can be viewed as basic (Battison, 1978). For Australian Sign Language (Auslan), which is closely related to BSL, it is proposed that the set of unmarked handshapes consists of, in addition to B, 5, A and G, the ‘O’ (a circle), ‘bc’ (an arc) and ‘S’ (like A, but with the thumb folded over the fingers) handshapes (Johnston and Schembri, 2007).

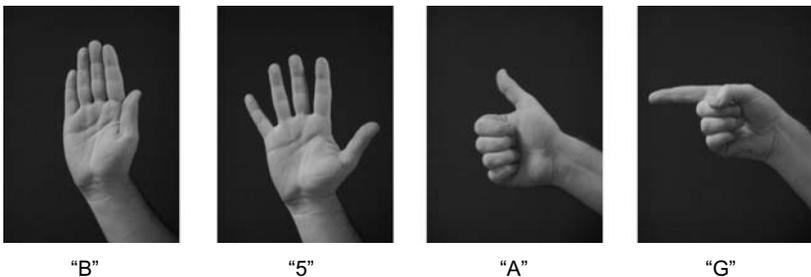


Figure 10.3a Unmarked handshapes B, 5, A, G

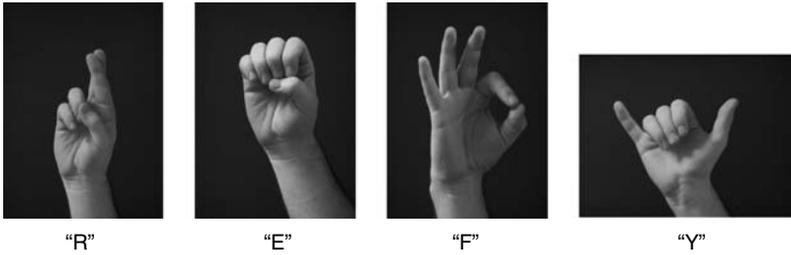


Figure 10.3b Marked handshapes R, E, F, Y

3.4 Language Contact

Just as spoken languages are influenced by contact with one another, so sign languages are influenced by surrounding spoken languages and by other signed languages. For example, although BSL is independent from English, it has been influenced by English. One of the things it has borrowed is the mouth patterns of English words. Some mouth patterns are easily identifiable as corresponding to English words, but others have been modified so that they cannot be identified by non-signing English speakers. Mouth patterns can serve to disambiguate two signs which would otherwise be identical, for example, BATTERY and UNCLE (Figures 10.4a and 10.4b).



Figure 10.4a BATTERY



Figure 10.4b UNCLE

The second area of major influence from English is fingerspelling, which is a way of spelling the letters of the alphabet on the hands. Figure 10.5a shows how the letters 'M', 'N' and 'C' are represented in the BSL manual alphabet. Names of people and places are frequently finger spelt, as are words for which there is no equivalent BSL sign. On occasion, finger-spelt forms become lexicalized, and as they are adopted into the BSL lexicon their phonology changes so that they obey the phonological constraints of BSL and become phonologically closer to signs in the core lexicon. For example, the English towns MANCHESTER and NEWCASTLE are not finger spelt in full, but are reduced to the letters M-C and N-C, respectively. Figure 10.5a shows that 'M', 'N' and 'C' require different numbers of selected fingers (3, 2 and 1 respectively, plus the thumb). However, in the lexicalized forms of these city names, the 'C' takes on the same number of fingers as the 'M' or 'N', so that the number of selected fingers remains the same throughout the sign – an example of assimilation (Figures 10.5b and 10.5c).



Figure 10.5a Letters 'M', 'N' and 'C' in the BSL manual alphabet



Figure 10.5b The 'C' in BSL MANCHESTER



Figure 10.5c The 'C' in BSL NEWCASTLE

One of the ways in which a sign language is influenced by contact with other sign languages is in its repertoire of handshapes. The hand can assume a large number of configurations. Just as spoken languages select different sets of possible speech sounds, so do sign languages choose different sets of possible handshapes. ASL and Irish Sign Language (ISL) have a large repertoire of handshapes, due in part to their having a one-handed finger-spelling system, and some of these handshapes are not generally seen in BSL signs. Due to contact between the British Deaf communities and those from the United States and Ireland, borrowings from ASL and ISL into BSL are not uncommon, and signs derived from the one-handed manual alphabets have proved particularly easy to adopt. For example, the 'R' handshape, with the index and middle fingers crossed (see Figure 10.3b), has only marginal status in BSL, occurring only in the sign for HOPE/WISH (the same form as the gesture for 'hope/wish' used in the hearing population; Brien, 1992), but it has been adopted in initial letter signs borrowed from ASL (e.g. ROCHESTER)⁸ and ISL (e.g. RELIGION). Similarly, the sign for Europe uses the 'E' handshape from the one-handed alphabet (see Figure 10.3b).

3.5 Models of Sign Structure

This section considers some important models that have been proposed for sign structure. We will see that sign-language phonologists also employ such terms as 'phoneme', 'segment', 'feature' and 'syllable', even though these clearly have different exponents than in spoken language.

Stokoe (1960) considered handshape, movement and location to be equivalent to the phonemes of spoken languages, except that unlike the phonemes of speech, they are organized simultaneously. Linguists have subsequently argued that signs do in fact contain sequential structure (e.g. Newkirk, 1998). In BSL MORNING, for example, the dominant hand makes contact with the contralateral side of the chest, then moves across to make contact with the ipsilateral side (Figure 10.1b). The handshape is identical throughout the sign, but the phonological representation needs to make reference to the starting point, the movement and the end point. Researchers are far from agreeing how this sequentiality should be represented in models of sign language structure, whilst also allowing for the representation of simultaneous aspects.

The first model to capture the sequential nature of signs also rejected Stokoe's classification of the major phonological categories. The Move-Hold Model (Figure 10.6) instead proposes two types of sequentially ordered segments (i.e. timing units): Movements (M), where the hand or hands move, and Holds (H) where they are held still (Liddell, 1984; Liddell and Johnson, 1989). Signs consist of sequences of Ms and Hs just as spoken words are made up of

sequences of vowels and consonants. Indeed, Liddell and Johnson proposed that Ms are analogous to vowels and Hs to consonants. Certainly, it is widely accepted that movements are like syllable nuclei in that they are essential for a sign to be well formed (Brentari, 1998; Perlmutter, 1992; Sandler, 1993; Wilbur, 1987).

Stokoe saw each handshape as an indivisible ‘phoneme’. However, subsequent work in spoken language phonology showed that breaking down phonemes into smaller units, or ‘features’, gives much greater explanatory power when accounting for phonemic inventories and phonological processes (see Botma et al., this volume). In the Move-Hold Model, Ms and Hs are associated with feature bundles, just as segments in spoken language are in SPE-type feature models (Chomsky and Halle, 1968). The Move-Hold Model also makes use of insights from Autosegmental Phonology (Goldsmith, 1976a). For example, as is the case in spoken language, features in sign language do not necessarily have a one-to-one relationship with the signs to which they are associated. This is shown, for example, in Figure 10.6 for the ASL sign LIKE (Liddell, 1990), where the M segment is characterized by the features of the Hs that precede and follow it.

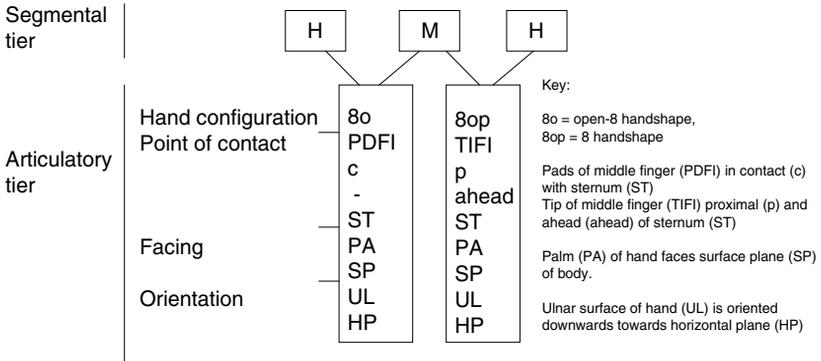


Figure 10.6 The ASL sign LIKE in the Move-Hold Model (from Liddell, 1990)

Despite its intuitive appeal, there are various problems with the Move-Hold Model (see Sandler and Lillo-Martin, 2006, for a detailed discussion). An obvious drawback is that by specifying all features on the H segments, even those that are shared across the sign (e.g. ST, PA, SP, UL and HP in LIKE in Figure 10.6), many features are represented redundantly. This results in the model predicting many types of signs which are actually unattested, including those that break the constraints outlined above.

The Hand Tier Model (Sandler, 1989) seeks to eliminate these redundancies. This model retains the sequentiality of the Move-Hold Model but revives handshape, movement and location as major phonological categories. The canonical form of a monomorphemic sign is shown in Figure 10.7. Movements (M) and Locations (L) are organized in sequence and are comparable in some ways to the Ms and Hs of the Move-Hold Model, but they are on a separate autosegmental tier to handshape and orientation, which are together termed Hand Configuration (HC).⁹

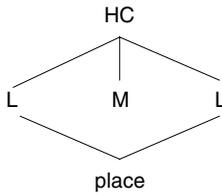


Figure 10.7 Canonical form of a monomorphemic sign in the Hand Tier Model (from Sandler and Lillo-Martin, 2006)

Feature geometry (Clements, 1985) has provided a valuable framework for elucidating hand configuration. Because of modality differences, the types of features in spoken and signed languages look very different: features such as [nasal] and [voice] obviously have no part to play in the latter. Instead, sign language features reflect the articulatory possibilities of the hands: for example, [all] indicates that all the fingers are selected, and [flex] indicates that the fingers are bent or curved.¹⁰ Different feature combinations are assumed, giving rise to different handshapes. An advantage of feature theory is that it makes it possible to capture markedness effects: marked handshapes have more complex representations in that they require more nodes and more features in their specification (see Sandler and Lillo-Martin, 2006, for discussion of these issues).

As in spoken languages, sign language features cluster into classes, motivated by the physical architecture of the articulators. Two such feature classes are 'selected fingers' (which includes the feature [all]) and 'finger position' (which includes [flex]). The selection of, for instance, all the fingers of the hand, or just one, is separate from the position of those fingers. In a sign where the handshape changes, it is only the position that changes, not the fingers that are selected (cf. the Selected Fingers Constraint, discussed in Section 3.2). This in turn motivates a feature hierarchy in which the position node is dominated by the selected fingers node (Sandler, 1989).

Also dominated by the selected fingers node in the Hand Tier Model is orientation. Evidence in favour of this comes from Newkirk et al.'s (1980)

aforementioned ‘slips of the hand’ study, where handshape substitutions included orientation as well (Sandler, 1989). Additional evidence for this comes from assimilation in compounds. For example, in ASL, *OVERSLEEP*, a compound of *SLEEP*^*SUNRISE*, either orientation or both handshape and orientation spread from *SUNRISE* to *SLEEP* (Sandler and Lillo-Martin, 2006). In this model, then, orientation is not a major category on a par with location, movement and handshape, as originally proposed by Battison (1978).

An alternative model of sign structure is the Prosodic Model (Brentari, 1998). Whereas the Hand Tier Model represents path movement as an M position on the segmental tier and internal movement as branching structure in the hand configuration category, the Prosodic Model represents both types of movements in the same branch of structure. The Prosodic Model separates movement, or ‘prosodic’, features from ‘inherent features’, that is, all features that characterize the whole sign and have no internal sequentiality. One reason why Brentari unifies path and internal movements is that each makes a sign well formed. Another is that for signs with both a hand-internal and a path movement, the internal movement is temporally linked with the beginning and end of the path movement; hence, the prosodic features of such signs are linked to units on the timing tier in the same manner.

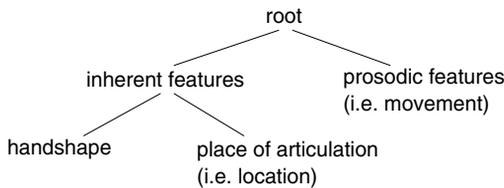


Figure 10.8 The Prosodic Model (from Brentari, 1998)

As the vast majority of signs have only one movement (or two simultaneous movements), the sign is usually monosyllabic (Coulter, 1982). This is reflected in the signs shown in Figures 10.6, 10.7 and 10.8. Despite the preference for monosyllabicity, sign languages can create a large number of phonologically distinct words because they have a larger set of features than spoken languages (see Sandler, 2008, for discussion of this). For example, whereas Halle (1992) proposes just 18 features for spoken languages, sign language phonologists have proposed anything from 30 (Sandler and Lillo-Martin, 2006) to 46 (Brentari, 1998).

Consider this last point in relation to the role of iconicity in sign language (see Section 2.2). Relevant to all models of sign structure is the question of

whether every single feature is necessarily part of the phonological feature inventory, or whether some features are more properly lexically specified because they are iconically motivated. Van der Kooij has argued for the latter position (van der Kooij, 2002). She argues that some features of handshapes and location are anomalous. For example, the location of KIDNEYS in Sign Language of the Netherlands (SLN) is the back, a location not found in any other signs of the language.

Such anomalous features are found in other languages too. For example, the BSL sign for the supermarket chain ASDA is borrowed from the advertisements for that store, whereby a shopper pats the change in her back trouser pocket; this location, too, is used in no other signs in the language. Similarly, the BSL sign for SCOTLAND is a double movement of the bent arm from the shoulder, against the side of the body, which looks quite unlike other signs, but is motivated by the movement involved in playing that traditional Scottish instrument, the bagpipes (Brien, 1992). Van der Kooij's claim is that the phonological feature inventory should not contain such rarely used, iconically motivated, features.

4. Case Study: Segmenting the Sign Stream

Word recognition in any language, whether spoken or signed, requires a continuous language stream to be segmented into discrete lexical units. Segmentation need not, of course, involve lexical access: it helps the listener or signer to identify word breaks, but recognition of the units bounded by those breaks is not a necessary step. Do spoken and signed languages use the same strategies for segmenting words, or does modality affect the segmentation process? One's *a priori* prediction would probably be that segmentation strategies would differ since, as we have seen, speech unfolds over time in a much more serial fashion than sign language does.

A group of psycholinguists led by Eleni Orfanidou set out to test whether a constraint proposed to guide segmentation in spoken language – the Possible-Word Constraint (PWC) (Norris et al., 1997) – also plays a role in the segmentation of sign languages (Orfanidou et al., 2010). According to the PWC, listeners make use of the knowledge that word boundaries should not be placed in such a way as to leave vowel-less 'residues'. For example, listeners have more problems perceiving 'see' in a nonsense word like 'seesh' than in 'seeshub'. This is arguably because breaking up 'sheesh' into 'see' and 'sh' leaves a residue ('sh') which, since it lacks a vowel, is not a possible English word. 'Shub' on the other hand does contain a vowel, and so is a possible English word. The PWC thus helps solve the segmentation problem by disfavouring lexical parses

that include impossible words. One could imagine that if signers also used the PWC they would benefit in the same way: they would avoid parses with impossible signs, and recognize new signs more easily. The question is whether this is what actually happens.

Orfanidou et al. investigated this possibility in BSL. Their experimental task was a sign-language variant of the word-spotting task (Cutler and Norris, 1988), which has provided the primary evidence for the PWC in a variety of spoken languages. In the word-spotting task, listeners hear a list of nonsense sequences, press a response button when they detect a real word embedded in the nonsense sequences, and say aloud the word they have spotted.

This task, and the manipulation of impossible- and possible-word contexts, were adapted for BSL. Deaf adult signers saw nonsense sequences consisting of two signs. On some trials, the second sign was a real BSL sign. The participants' task was to press a button when they spotted a BSL sign, and then sign it to a camera. Targets appeared after a nonsense sign that was either possible or impossible in BSL. There is, however, a fundamental difference in what can constitute an impossible word between signed and spoken languages. Because every BSL sign must contain a location, handshape and movement – it is impossible to produce a sign without these parameters – it is not possible to leave out material in the way that a vowel can be left out in a spoken nonsense sequence. Orfanidou et al. therefore created impossible nonsense signs by adding superfluous phonological material to existing signs. For example, disyllabic signs with straight + arc movement, or one-syllable signs with both a handshape and an orientation change, are not permissible (Uyechi, 1996).¹¹

Orfanidou et al. hypothesized that Deaf signers would be faster and more accurate in identifying signs that are preceded by a possible nonsense sign. Three groups of adults aged between 18 and 60 were tested: native signers (exposed to BSL before 5 years of age), childhood learners of BSL (exposed to BSL between 6 and 12 years of age) and adolescent learners. As predicted, participants were faster and made fewer errors in detecting real BSL signs in nonsense contexts that were possible signs than in contexts that were impossible signs. There was no effect of age of acquisition of BSL – the native signers, for example, did not perform faster or more accurately than the later learners. The results from all groups are consistent with the results of word-spotting tasks in languages such as English: sign spotting is easier when the sign is embedded in a phonotactically possible context, as predicted by the PWC.

At this point one methodological concern should be noted. In Orfanidou et al.'s experiment impossible-sign contexts were created in a different way than in the spoken language versions of this task. The impossible-sign contexts were created by using superfluous and illegal combinations of phonological parameters. This raises the question whether the disadvantage for the impossible-sign

context condition reflects a dispreference for contexts with illegal phonotactics, rather than a dispreference for a lexically nonviable sign. Because impossible signs can only be created by making them phonotactically illegal, there appears to be no way to avoid this problem. However, it is possible to address this issue using speech.

To do this, Orfanidou et al. ran a second experiment, this time using native speakers of Dutch. They created three types of nonsense contexts in which real words were embedded. If, for example, the target word was *long* ('lung'), one of the nonsense contexts, *slong*, left a single consonant which could not be a word, the second, *schruuslong*, left a syllable that had a legal onset cluster, and the third, *sfruuslong*, left a syllable that had an illegal onset cluster. This third context was therefore analogous to the impossible context used in BSL. The prediction was that Dutch listeners would find it harder to detect words in the single consonant context, but not in the two syllabic contexts. If, on the other hand, performance in the illegal syllabic contexts was found to be poorer than in the legal syllable contexts, then this would suggest that phonotactically illegal material in a nonsense context is dispreferred in segmentation, and hence undermine Orfanidou et al.'s PWC account of the BSL results.

In their follow-up experiment, Orfanidou et al. found an effect of context for both reaction times and accuracy. Words in single consonant contexts were spotted more slowly and less accurately than words in syllabic contexts, but there were no differences between the two types of syllable contexts. The results therefore indicate that the parallel effects across modalities are likely to be due to the same segmentation algorithm – the Possible Word Constraint – even though there are differences in what can constitute an impossible word in sign and speech.

5. Conclusion: Current Controversies in Sign Language Phonology

Some of the issues that are currently being debated in sign language phonology are specific to the field itself. Others impinge on broader issues, and are therefore of interest not just to sign language phonologists but to phonologists (and linguists) in general.

One of the most interesting linguistic conundrums that sign languages may be able to solve is the puzzle of how phonology evolved. Although all known sign languages are relatively young, some, such as ASL, BSL and French Sign Language, are hundreds of years old, typically dating from the gathering together of deaf children in the first residential schools for the deaf. In contrast to those more formal, school-based and urban sign languages are languages that have arisen in close-knit village communities, where consanguineous marriages have resulted in a high incidence of genetic deafness. Examples of

such village sign languages include Adamorobe Sign Language in Ghana (Nyst, 2007), Yucatec Mayan Sign Language in Mexico (Johnson, 1991) and Al-Sayyid Bedouin Sign Language (ABSL) in Israel (Sandler et al., 2005). These sign languages have arisen independently of their respective national sign languages, and they allow us to study the evolution of language in progress.

According to Hockett (1960: 95), duality of patterning is a basic design feature of human language, but evolved relatively late

because one can find little if any reason why a communicative system should have this property unless it is highly complicated. If a vocal-auditory system comes to have a larger and larger number of distinct meaningful elements, those elements inevitably come to be more and more similar to one another in sound. There is a practical limit . . . to the number of distinct stimuli that can be discriminated, especially when the discriminations typically have to be made in noisy conditions.

Almost 40 years after Stokoe's discovery of duality of patterning in ASL, Aronoff et al. have made the striking claim that ABSL lacks phonology (see Aronoff, 2007; Aronoff et al., 2008). As evidence for their claim, they observe that although ABSL has a conventionalized vocabulary, there is considerable variation in form. For the sign BANANA, for example, both of the signers Aronoff et al. investigated produced a movement indicating the peeling of a banana located at chest height. However, one signer used an 'F' handshape (see Figure 10.3b) while the other used a 'B' handshape (see Figure 10.3a). Although these two handshapes are contrastive in Israeli Sign Language, as indeed they are in ASL and BSL, they do not appear to be used contrastively here. This leads Aronoff et al. to conclude that 'signs tend to be exemplified by a set of tokens centred around a prototype, where each token of a given prototype may have a different handshape, location or movement, but conveys the same concept' (2008: 137). They go on to argue that ABSL has been able to develop into a full-fledged linguistic system without the benefit of phonology because of signing's visual medium, which allows for direct iconicity.

It may be that ABSL will develop phonology as it ages, perhaps simply as a function of the size of its vocabulary, as suggested by Hockett (and as modelled by Nowak and Krakauer, 1999). Indeed, Aronoff and colleagues present fascinating new data showing that phonological and lexical regularity may start within the phonological system of a family unit, and may be emerging in the youngest signers (Sandler and Aronoff, 2007). As such, ABSL and other village sign languages therefore provide a window onto the evolution of phonology, a window that would not be open to us if phonologists studied only spoken languages.

6. Acknowledgements

I am indebted to Gary Morgan, whose paper on the acquisition of BSL phonology at the 2004 Manchester Phonology Meeting inspired me to learn and research BSL. I owe pretty much all I know about sign language phonology to him and to our colleagues at UCL's Deafness, Cognition and Language Research Centre and City University's Sign Language and Gesture Laboratory. In particular, I thank Adam Schembri and Kearsy Cormier, who teach an enthralling sign language linguistics class at UCL, and Eleni Orfanidou, who shared the results of her ingenious sign segmentation experiments ahead of publication. Neil Fox and Tanya Denmark posed for the photographs, which Annika Pabsch took. Adam Schembri, Gary Morgan, Rachael-Anne Knight, Annika Pabsch and Robin Thompson commented on earlier drafts and saved me from making too many errors. Three anonymous reviewers and the editors also gave very useful feedback, for which I am grateful. Needless to say, I take full responsibility for any errors that remain. Finally, I would like to acknowledge the Leverhulme Trust, whose award of an Early Career Fellowship supported the writing of this chapter.

7. Notes

1. ASL and BSL have different origins. ASL is actually related to French Sign Language (LSF): a French teacher of the Deaf, Laurent Clerc, introduced LSF to the United States when he helped establish the first school for the Deaf in 1817. BSL has a long history, but because sign languages leave no written trace, it is difficult to pinpoint with any accuracy when it originated. However, in 1666 the English diarist Samuel Pepys recorded an encounter in London with a deaf boy who used sign language.
2. In the sign language literature, deaf with a lowercase 'd' is used to refer to the audiological condition of deafness, whereas Deaf with an uppercase 'D' refers to deaf people who are part of a community of sign language users.
3. The gloss for a sign is conventionally written in capital letters.
4. Sine-wave speech tokens contain the same temporal and prosodic information as speech at the same frequencies, but lack the distinctly biological quality of true speech.
5. The 'A' handshape and the head location described in these examples are probably best analysed not as morphemes, but rather as examples of form symbolism, akin to examples of sound symbolism in spoken languages (Brennan, 1990). Form symbolism is a type of iconicity, which, as we have seen, is more common in signed than in spoken languages.
6. These constraints have been proposed for ASL. How universal they are, particularly for non-Western sign languages, requires further research.
7. The convention is for handshapes to be named after the letters they represent in the ASL alphabet or counting system.
8. The National Technical Institute for the Deaf in Rochester is the first and largest technological college in the world for students who are deaf or hard of hearing.

9. In this model, the place category is associated only with L because place features belong to the location category (see Sandler and Lillo-Martin, 2006). Phonetically, place spreads to characterize the movement.
10. In the Hand Tier Model features are unary, contrary to what is assumed in traditional generative phonology, where features are binary (e.g. Chomsky and Halle, 1968). There is no evidence that the features used to specify handshape have minus values (see Sandler and Lillo-Martin, 2006).
11. Uyechi's claims were made for ASL, but in the absence of evidence to the contrary, Orfanidou et al. hypothesize that they also hold for BSL.

11

Phonological Acquisition

Nivedita Mani

Chapter Overview

1. Introduction	278
2. Acquiring the Sound Patterns of Language	280
3. Phonological Specification of Lexical Entries	288
4. Summary	295
5. Notes	297

1. Introduction

Parents and caregivers of infants are often curious about the timeline of their child's linguistic entry into society. Most language acquisition researchers will have been faced with the question, 'When does language acquisition begin?' This curiosity is legitimate – typically, infants do not utter their first words until the second year of life, so there is little overt demonstration of language use. It is only between 6 and 10 months of age that infants reliably demonstrate comprehension of parental input (i.e. words) in experimental circumstances (Bortfeld et al., 2005; Tincoff and Jusczyk, 1999). For instance, Tincoff and Jusczyk (1999) found that 6 month olds correctly associated videos of their mother and father with the words 'mommy' and 'daddy', respectively. Given that such laboratory precision, as in Tincoff and Jusczyk (1999), is not available to the average parent, one cannot be faulted for holding the assumption that the 1-year-old is pre-linguistic.

Contrary to such expectations, language experience, that is, hearing spoken language, begins to shape infant speech perception remarkably early in life. Four-day-old infants successfully discriminate their native language from an unfamiliar language (Mehler et al., 1988) and also discriminate two unfamiliar, that is, non-native languages (Mehler and Christophe, 1995; Nazzi et al., 1998). Given that discrimination performance was unrestricted by infants' familiarity with the languages, however, success was probably driven by salient acoustic

properties of the languages tested, that is, their prosodic characteristics. Indeed, Nazzi et al. (1998) found that while French neonates could discriminate between English and Japanese, they could not discriminate between English and Dutch sentences. Failure was attributed to the prosodic differences between the pairs of languages – Japanese is mora-timed, while both English and Dutch are stress-timed languages (Mehler et al., 1988; Nazzi et al., 1998). Interestingly, infants' precocious ability to discriminate unfamiliar languages disappears by 2 months of age – a finding explained by suggesting that older infants may already have narrowed their perceptual focus to utterances that share characteristics with their maternal language, while ignoring those that do not (Mehler et al., 1988).

Some researchers place the onset of speech perception earlier still, based on findings that fetuses show a preference for familiar texts (read to them *in utero* by their mothers) over unfamiliar texts (DeCasper et al., 1994). Similarly, both fetuses (Kisilevsky et al., 2003) and new-borns (DeCasper and Fifer, 1980) show a preference for listening to their mother's voice over a female stranger's voice. Such a familiarity preference suggests that fetuses and new-borns alike encode the perceptual characteristics of the speech stimuli they have been exposed to. This encoding might mark the beginnings of linguistic identity and indicates a high level of processing at an early age, since the heard stimuli need to be compared to previously registered stimuli for successful performance. These initial findings have been taken to suggest that infants are born with a 'bias' for listening to speech (Peña et al., 2003; Vouloumanos and Werker, 2007), with argument abounding about whether this bias for speech is experience-dependent (given the number of studies showing *in utero* familiarization with speech, Kisilevsky et al., 2003) or possibly 'tantalizing[ly]' (Vouloumanos and Werker, 2007: 163) experience-independent (on the basis that a similar bias is not found in children with Autism Spectrum Disorder: Klin, 1991; but see Rosen and Iverson, 2007).

Such precocity is not restricted to infant speech perception but is also displayed in infants' early vocalizations – termed protophones or precursors to speech (Oller, 1980; Oller et al., 1999). These vocalizations typically evolve from quasivowels at 2 months of age (sounds sharing some features with later vowel productions), to cooing between 2 and 3 months of age (quasivowels that include movement of articulators during vocalization), to vowel-like sounds between 4 and 6 months of age (includes greater use of the vocal tract to produce contrastive vowel and consonant-like sounds, also called marginal babbling) to canonical babbling at around 6 months of age. Canonical babbling marks the onset of syllable-like productions with fast movements from consonant-like to vowel-like sounds, often duplicating syllables, for example, *mama*, *dada*. However, these syllables are not counted as words because we cannot assume that infants assign meaning to these syllables (but see Borfeld et al., 2005; Tincoff and Juszcyk, 1999).

Some researchers suggest that, by the end of the first year of life (between 6 and 10 months), babbling begins to take on the characteristics of the infants' native language ('babbling drift': Brown, 1958; Engstrand et al., 2003), with similarities in the prosodic characteristics (fundamental frequency of vocalic sounds and stress patterns of the language) and the frequency of consonant sounds in the infants' babbling and their native language (Boysson-Bardies et al., 1984, 1986, 1989, 1992). However, others, arguing for independence between language-specific exposure and early vocalizations, report not finding a similar language-specific effect on infants' babbling (Atkinson et al., 1969; Olney and Scholnick, 1976; Thevenin et al., 1985; Oller et al., 1997¹). The ultimate solution to this contradiction may lie in further analysis of the individual differences between infants, the putative asynchrony between the prosodic and segmental characteristics of infant and adult babbling (pers. comm. between Lindblom and Engstrand reported in Engstrand et al., 2003)², or the salience of the specific linguistic characteristics under investigation (Locke, 1993).

These studies suggest that the infant makes an early start on language learning. The aim of this chapter is to detail these important landmarks of infant language acquisition up to and including the second year of life, with specific focus on phonological development – infants' acquisition of the sound patterns of their native language (Section 2), and the use of these sound patterns in encoding the words they subsequently acquire (Section 3). This discussion will focus not just on identifying the onset of different kinds of linguistic knowledge, but also the factors that influence this development and, thereby, contribute to the shaping of phonological identity. Given the focus on production data in the chapter on 'Methodology in Phonological Acquisition' (Zamuner and Johnson, this volume) and the volume of data in the field of infant speech perception, this chapter will prioritize data on infant speech perception in phonological development, and the implications thereof.

2. Acquiring the Sound Patterns of Language

Lieberman et al. (1961) report that adults perform better at discriminating sounds that straddle phonemic boundaries (e.g. /p/ vs /b/) compared to sounds that are contained within the same phonemic category (e.g. two tokens of /p/ that vary in voice onset time), despite maintaining the same acoustic difference across the two contrasts (i.e. 5 milliseconds lag in voice onset time (VOT)). Even more surprising, then, was the finding that 1 month olds discriminate sounds that vary along a VOT continuum in a manner similar to adults – with greater sensitivity to between-category differences (i.e. /p/ versus /b/) than within-category differences (*categorical perception*: Eimas et al., 1971). Such categorical perception persists even when the sounds are not native to their language

(Kikuyu infants tested on the English voicing distinction not present in Kikuyu: Streeter, 1976). One obvious implication of this early ability to discriminate native and non-native sound contrasts would be that infants demonstrate a linguistic sensitivity to sounds from all languages, with pre-specification of the universal sound patterns of language.

However, this explanation has been overruled by decades of research showing that infants show similar categorical perception of non-speech sounds (Cutting and Rosner, 1974; Cutting et al., 1976; Jusczyk et al., 1980); whilst chinchillas demonstrate categorical perception of speech sounds (Kuhl and Miller, 1975, 1978). Given that one cannot credit chinchillas with linguistic sensitivity to speech sounds or human infants with pre-specification of the characteristics of non-speech sounds, the argument that infants' initial sensitivity to phonemic boundaries is linguistic or speech-specific loses weight.

As with the studies on language discrimination reported above (Mehler et al., 1988; Mehler and Christophe, 1995), this early language-universal precocity does not last long. Seminal work in this regard goes back to Werker and Tees' research on the Hindi dental and retroflex contrast /ta/-tâ/, the dental voiced aspirated and the voiceless aspirated contrast /t^ha/-d^ha/, and the Salish glottalized velar and uvular contrast /ki/-çqi/ (Werker and Tees, 1983, 1984). Similar to the studies conducted by Streeter (1976), Werker and Tees report that English 6 month olds successfully discriminate all three contrasts, despite the fact that none of these contrasts are native to the English language. However, older infants (between 10 and 12 months of age) lose their ability to discriminate these non-native contrasts, a loss in sensitivity that is maintained throughout their childhood (Werker and Tees, 1983) and adulthood (Werker et al., 1981). Importantly, language experience does not result in changes to non-native sensitivity alone. The end of the first year of life also heralds changes in infants' discrimination of native contrasts. For instance, Kuhl et al. (2006) show that infants become *better* at discriminating native language contrasts between 6 and 12 months of age (see Rivera-Gaxiola et al., 2005 for similar results).

Researchers have proposed a number of factors to explain this developmental change between 6 and 12 months of age, and in what follows, we consider each of these factors in turn. Invariably, these factors relate to the quality and quantity of language exposure the infants experience. Even selection models (Eimas, 1975; Liberman et al., 1967; Liberman and Mattingly, 1985), arguing for innate pre-specification of phonetic units, allow a role for language experience – infants' failure to discriminate non-native contrasts is due to the loss of adequate specification of these contrasts because of their absence in the infants' immediate native language environment. However, selection models have, since, become unpopular because of findings that infants successfully discriminate non-speech sounds without pre-specification of the characteristics distinguishing these sounds (Cutting and Rosner, 1974; Cutting et al., 1976),

and results suggesting that adults can discriminate some non-native speech sounds, given sufficient language-specific training (Pisoni et al., 1982; Werker and Logan, 1985; Rivera-Gaxiola et al., 2000). What characteristics of language exposure, then, drive such drastic changes in infant sensitivity?

2.1 Acoustic Information

One explanation for the early universal discrimination abilities reported above implicates the acoustic characteristics of the stimuli presented to infants and adults – specifically, the temporal characteristics of these stimuli (Juszyk et al., 1980). Hirsh (1959) and Hirsh and Sherrick (1961) report that adults cannot accurately specify the order of presentation of two sounds when the onset of the two sounds were separated by less than 20 milliseconds (the voice onset lag that distinguishes voiced and voiceless stops). These findings suggest that the temporal limitation that underlies categorical perception might reflect sensitivity to salient acoustic properties of signals (speech or non-speech), such as the onset lag between stimulus events, that is, between onset of vocal fold vibration and stop release (Juszyk et al., 1980) or, as has been argued by Stevens and Klatt (1974) the presence or absence of F1 transition (but see Lisker, 1975).

Evidence supporting an important role for acoustic salience in speech sound discrimination comes from experiments suggesting that early universal discrimination sensitivity does not, as previously thought, stretch to all speech sounds – Narayan et al. (2010) report that English 4 month olds do not discriminate acoustically non-salient non-native Filipino contrasts (/na/ - /ŋa/). Furthermore, Filipino 6 month olds also do not discriminate this native contrast, arguing even more persuasively for a role for acoustic salience in early performance. Only between 10 and 12 months of age do Filipino infants successfully discriminate the same acoustically non-salient but native contrast. Early discrimination abilities might, therefore, be attributed to the specific acoustic characteristics of the contrasts presented to infants. In contrast, later language-specific discrimination might rely on the recently established phonemic categories in the infants' native language, fine-tuning the infants' perceptual system into ignoring acoustic variance irrelevant to their native phonemic categories (Kuhl, 1991, 2000, 2008).

2.2 Frequency

One can easily imagine a role for the frequency of different phonemes in the native language environment in driving infants' sensitivity to different

phonemic contrasts, that is, not only must the phonemes be present in the input, but infant sensitivity can also be shaped by the frequency of one or both members of the phonemic contrast tested.

Infants are adept at attending to the statistical frequency of phoneme clusters in their native language and using this to selectively direct their attention to their language input (Jusczyk et al., 1994; Saffran et al., 1996). Indeed, the literature on frequency effects in language development has predominantly focused on phoneme clusters, rather than the frequency of individual phonemes. One study that focuses attention on the frequency of individual phonemes compares infants' sensitivity to contrasts less or more frequent in their native language: infants perform better in their discrimination of the non-native dorsal contrast than a non-native coronal contrast (Anderson et al., 2003). Since infants are exposed to almost three times as many coronals (157297) than dorsal sounds (55720) in American English child directed speech, the higher frequency of the coronals may have led to the infants acquiring the native language [coronal] feature earlier. In keeping with the perceptual fine-tuning outlined above, the earlier instantiation of the coronal feature might lead infants to ignore acoustic variance contributing to coronal contrasts not present in their native language and absorbing the presented tokens into their native language coronal categories.

2.3 Distribution of Phonemic Inventory

Whilst it might seem strange to distinguish between the distribution and the frequency of the phonemes in the infants' native language input, there are crucial differences between the two factors. One can view the distribution as the mere presence or absence of a phoneme in the infants' environment, while frequency calls more on the number of times a phoneme is presented to the infant. Unsurprisingly, a number of studies have found that familiarizing infants with different distributions of phonemes has a robust influence on their subsequent discrimination of similar sounds.

Maye et al. (2002) familiarized English infants with a sequence of sounds varying in voice onset time along the /da-/ta/ continuum. One group of infants was familiarized with a bimodal distribution such that they were presented with stimuli along the endpoints of the continuum (prototypical /da/ and /ta/ sounds), while another were presented with a unimodal distribution of stimuli from the middle of the continuum (ambiguous between /da/ and /ta/). Despite the presence of the /da-/ta/ contrast in their native language, exposure to the unimodal distribution reduced infants' discrimination of the contrast. Similarly, Maye et al. (2008) found that exposure to a bimodal distribution of

phonemes improved 8 month olds' sensitivity to a contrast that they typically find difficult to discriminate, namely pre-voiced versus short lag stop consonants (Aslin et al., 1981). Furthermore, exposure to a bimodal distribution allowed infants to generalize the features of the phonemic repertoire and extend this information to other as yet unfamiliar sounds – exposure to the bimodal distribution of one stop contrast (e.g. /da/-/ta/ with a pre-voicing versus short lag difference) improved infants' performance on an unfamiliar contrast (e.g. /ka-/ga/ with a similar pre-voicing versus short lag difference).

Furthermore, McMurray and Aslin (2005) report that exposure to a unimodal distribution may have the effect of improving infants' sensitivity to within-category distinctions. Two groups of 8 month olds were exposed to lists of words that began with either /b/ or /p/. Following familiarization, infants displayed greater sensitivity to within-category variations of the sound (either /b/ or /p/) that they had been exposed to relative to between-category variation (/b/ vs /p/).

2.4 Timing of Exposure

Language research since the 1960s has been attracted to Lenneberg's (1967) doctrine on the role of age and maturation on language acquisition. The notion that a language learner cannot attain native-like fluency in language post-puberty is a lore recognized by researchers and laymen alike. More recently, this account has been the focus of attention in the realm of phonological development. The results of Werker and Tees (1983) suggest that sensitivity to unfamiliar phonemic contrasts is lost by the end of the first year of life. However, what happens to contrasts that adults were exposed to as infants, but never since? Pallier and colleagues examine this issue with recourse to a group of Korean subjects adopted by French-speaking families (between 3 and 8 years of age). As adults, they showed poor discrimination of Korean voiceless consonants (not native to French) irrespective of whether they had recently been back to visit Korea or not (Ventureyra et al., 2004). In contrast, Oh et al. (2003) report a graded sensitivity to contrasts based on the amount of exposure Korean speakers have with their language – those without any exposure post childhood performed the worst in discriminating Korean contrasts, compared to those with limited exposure (few hours a month) post childhood. Finally, Højen and Flege (2006) report that although early Spanish learners of English successfully discriminate English vowels, there were still differences in the performance of early learners and English monolinguals. Taken together, these results suggest that whilst early exposure alone is not adequate to maintain sensitivity to non-frequent or non-native contrasts, adults with limited early exposure are advantaged over adult L2 learners.

2.5 Theoretical Perspectives

Researchers have proposed different approaches to theorize the influence of language experience and the development of native language phonology on non-native contrast discrimination. One school of thought, *Neural Commitment* (Native Language magnets: Kuhl, 1991, 1993, 2000, 2008), takes off from research establishing the influence of acoustic and distributional information on phonological development. Kuhl argues that infants use information regarding the distributional frequency of different phonetic segments in speech to map the locations of ‘modal values’ of phonemic categories (or phonemic prototypes). Very young infants (between 1 and 6 months) have not mapped their phonemic prototypes yet and use general auditory mechanisms to discriminate speech sounds (and non-speech sounds), hence their early ability to discriminate native and non-native acoustically salient contrasts. In acquiring language, infants become neurally committed to the sound patterns of their native language, distorting their perception of speech sounds in favour of their native language (see also Dietrich et al., 2007). Kuhl et al. (1992) demonstrate how such neural commitment might work. Once the native language prototypes are formed, these act as perceptual magnets for tokens of native language phonemes – By 6 months, American infants cannot discriminate between different tokens of the American vowel [i], while Swedish infants can (the vowel [i] is not native to Swedish). The reverse is also true: American infants can discriminate between different tokens of the Swedish vowel [y] while Swedish infants cannot. Thus, by 6 months of age, native language vocalic prototypes are in place.

Language experience and perceptual distortion also impact non-native speech sound discrimination. Iverson et al. (2003) suggest that the distributional properties of native language sounds warp the listeners’ perceptual space so as to make acoustic differences that are pertinent to discriminating different sounds in their native language more salient than non-relevant acoustic characteristics. This has the effect of rendering some non-native contrasts (whose acoustic differences are non-salient in the listeners’ native language) non-discriminable, while others (whose acoustic differences are required to differentiate native language phonemes) can be easily discriminated.

Arguing against a purely acoustic approach, Best and colleagues (Best, 1994; Best et al., 1988; Best and McRoberts, 2003) suggest that defining the threshold of acoustic salience that renders a contrast (native or non-native) discriminable or not may prove difficult, if not subjective. In contrast, using the notion of *Perceptual Assimilation*, Best suggests that young infants become attuned to gestural constellations (clusters of articulatory gestures) that fit their native language phonological repertoire, and use these constellations to discriminate sounds presented to them – that is, they do not attend to adult phonological

patterns, per se. Adults, on the other hand, pick up on phonological patterns (or similarities) and use this knowledge to assimilate the sounds of a non-native contrast to their native language sounds. This mapping onto native language phonemes takes place in terms of the articulatory characteristics of the non-native sound. By combining elements of Goldstein's *Articulatory Organ* hypothesis (Studdert-Kennedy and Goldstein, 2003), Best suggests that non-native phonemes that share the same primary articulator should prove harder to discriminate than phonemes that have different primary articulators (Best and McRoberts, 2003). Discriminability of non-native phonemes, therefore, rests on the mapping of these phonemes onto separate native language phonemes (Best et al., 2001). Non-native phonemes that map onto two separate native phonemes (Two-Category Discrimination) would prove easier to discriminate than non-native phonemes that map onto the same native phoneme (Single-Category Discrimination). There are exceptions to the non-discriminability of single-category contrasts though – if the articulatory characteristics of a non-native phoneme map onto a prototypical phoneme of a non-native contrast, while the other phoneme of the pair maps onto a non-prototypical token of the same phoneme, then the non-native contrast can be discriminated on the basis of the goodness of fit of the two non-native phonemes onto the native language phoneme (Category Goodness Discrimination).

In other words, the developmental change from language-universal to language-specific discrimination sensitivity is caused by the development of native language phonemic categories. After this period, discriminability of non-native phonemes rests on the articulatory mapping between the non-native sounds and the native phonemic categories. Adults should still be able to discriminate some non-native contrasts that map onto sufficiently distinct native phonemes, whilst failing to discriminate non-native contrasts that do not.

A final noteworthy contender to this list accords responsibility to domain-general cognitive development (Lalonde and Werker, 1995). These authors suggest that any account of phonological development brokering a structural change in phonological processing cannot explain how adults can, given sufficient training, regain a 'lost' sensitivity to non-native contrasts. They argue that change in performance must be related to improvement in post-perceptual domain general or cognitive abilities. These domain-general abilities act on the output of speech perception. Adding or reducing the cognitive load on processing can lead to changes in discrimination or speech perception abilities, without signalling a change in the representations or processes underlying speech perception. Werker and Lalonde report a correlation between infants' non-native speech perception abilities and their performance in domain-general cognitive tasks, suggesting an influence of cognitive development on

infant speech perception. The non-linguistic tasks included testing infants' visual categorization performance (Younger and Cohen, 1986 – to provide support for the hypothesis that an improvement in general categorization abilities could lead to better phonemic representation) and an object-search test (Piaget, 1954 – to test for a correlation between generic cognitive development and phonological development).

An important addition to this set of studies was recently completed by Conboy et al. (2008) – Conboy et al., report that 11 month olds' native language speech discrimination abilities were correlated with receptive vocabulary size, but not with infants' cognitive control scores. In contrast, infants' non-native speech discrimination abilities were negatively correlated with cognitive control scores (similar to Lalonde and Werker, 1995), but not with receptive vocabulary size. Conboy et al., use these results to argue for a role for cognitive development in speech perception, but only with regard to ignoring irrelevant non-native speech sound variation.

The emergent picture from this series of studies suggests a role for articulatory, acoustic and domain-general cognitive faculties in non-native speech sound discrimination (or the decline thereof) and acoustic properties and distributional information in native speech sound learning (and discrimination). This is not to suggest that these factors do not have an influence overall on phonological development, but rather to highlight the areas where there is an attested relationship between the specific faculties of speech perception and the factors outlined above. What becomes increasingly clear from this review is that no theory of phonological development is complete without paying adequate attention to the different factors outlined here, and a specific outline of their potentially separable influence on native and non-native speech perception.

By all accounts, the average 1-year-old has more than adequately categorized the sound patterns in her phonological repertoire. By 6 months of age, infants differentially categorize native and non-native vowel sounds (Kuhl et al., 1992), suggesting separable processing of native and non-native vowels. A similar differentiation of native and non-native consonantal patterns is in place by 10 to 12 months of age (Werker and Tees, 1983). By this age, infants also show sensitivity to the groupings of phonemic categories permissible in their native language (native language phonotactics: Jusczyk et al., 1993) – American 9 month olds listen longer to words that follow the phonotactic rules of English than words that do not (e.g. Dutch words allowing word-initial consonant clusters such as /zw/) – suggesting sensitivity to word-level groupings (or at the very least, suprasegmental groupings). The cogs appear securely in place for infants to begin representing their lexical repertoire in phonological terms towards the end of their first year of life.

3. Phonological Specification of Lexical Entries

Given the attested stability of infant phonological categories by the end of the first year of life, logic dictates that the words infants acquire should be represented in maximal phonetic detail to ease subsequent recognition and production of these words. Indeed, unsurprisingly, 7.5 months olds pay attention to phonetic detail in encoding familiar words (Juszyk and Aslin, 1995). Infants in this study were initially familiarized with a word, and then tested on their recognition of this word in unfamiliar sentences, and vice versa. Infants displayed recognition of the familiarized words only when the words were accurately pronounced, but not when the words were mispronounced by even a single segment (e.g. *cup* mispronounced as *tup*). Infants' sensitivity to mispronunciations of the familiarized word, *cup*, lends itself to the suggestion that infants' representation of the word was phonologically detailed, that is, fully specified with regard to the phonemes of the word.

One caveat to this finding is that we have no way of determining the lexical status of *cup* in the infants' mind when performing this task. Is *cup* here anymore a word than the sequence of phonemes 'teeb', had infants been familiarized with the words 'pretty baby'? We would not want infants to be similarly sensitive to variations in the phonemic structure of *cup* and *teeb*, thereby resulting in a similar response to *tup* (a mispronunciation of *cup*) and *teed* (in, for example, *pretty dog*). In other words, is infants' sensitivity to the mispronunciation of *cup* a demonstration of the level of phonemic detail associated with the lexical representation of this word, or a demonstration of infants' sensitivity to the familiarized grouping of the phonemes in *cup*? Given our knowledge of the excellence with which infants pay attention to statistical or distributional information in the speech stream (Saffran et al., 1996; Maye et al., 2002), we may not be able to differentiate between these two explanations in the current task (or indeed, any task that does not tap into the lexical status of the word).

Stager and Werker (1997), in contrast, use a word-object association task in an attempt to access the lexical status of phonemic clusters – infants were familiarized with the word *bih* in the presence of an unfamiliar object, and the word *dih* in the presence of yet another unfamiliar object. Following infants' habituation to these pairings, infants were tested on their sensitivity to the association between the novel 'words' and the novel objects. Contrary to expectations (raised by the Juszyk and Aslin study), 14 month olds did not discriminate incorrect pairings from correct pairings – infants did not think that *dih* was an inappropriate label for the *bih*-object, or vice versa. Two obvious solutions raise themselves in response to this finding: (i) infants are unable to discriminate between phonetically similar sounds, irrespective of objects associated with these sounds; (ii) infants are unable to associate words

with objects. Subsequent experiments ruled out both these explanations by showing that (i) infants are able to discriminate these sounds in the absence of a word-object association, and (ii) infants are able to correctly associate words with objects when the two words are more distinct phonologically (e.g. *lif* and *neem*).

The authors reach the conclusion that infants seem unable to discriminate phonologically similar words when these words are presented in the context of an object association, while displaying a keen sensitivity to phonological dissimilarities otherwise. Why would infants not display sensitivity to phonological mismatch only in the presence of an object association? There are two answers to this question, both of which rely on the assumption that the object association triggers the lexical representation of the phoneme cluster, thereby endowing it with lexical status (unlike in the Jusczyk and Aslin study).

One obvious explanation is that words in the infant lexicon are not encoded with sufficient phonological detail. Despite their noteworthy phonological repertoire, infants and young children do not use this phonological knowledge when it comes to encoding words in their native language – words are represented holistically or in a global fashion without adequate regard to segmental or featural level detail. Word recognition is conducted through matching of acoustic-phonetic information at a less detailed level of analysis, such that *dih* may sometimes be recognized as *bih* (Charles-Luce and Luce, 1990).

A second explanation for the above findings is that, despite sufficient detail in infants' representations of words, there may be cognitive limitations on infants' processing of sounds in word recognition (Stager and Werker, 1997) or attentional limitations on infants' processing of featural or temporal information in words (Charles-Luce and Luce, 1990). Thus, infants should not display sensitivity to mispronunciations of words early on, despite the lexical representations of these words being phonologically well specified.

Both explanations are challenged to varying degrees by the results of numerous studies examining infants' sensitivity to mispronunciations of the phonemic segments in familiar (Swingley and Aslin, 2000, 2002; Bailey and Plunkett, 2002; Ballem and Plunkett, 2005; Mani and Plunkett, 2007; Mani et al., 2008) and newly learnt words (Ballem and Plunkett, 2005; Mani and Plunkett, 2008a). Contrary to the hypotheses suggested above, these studies report that 14 month olds are sensitive to mispronunciations of the phonemic segments in words. As with the Jusczyk and Aslin (1995) study, sensitivity to a mispronunciation suggests that infants detect the mismatch between a heard word and their representation of this word. Their representation of this word must, therefore, contain adequate phonological detail to allow infants to detect such a mismatch. These findings provide support for the alternative view that, not only do infants have phonologically well-specified representations of words,

but also there appear to be few restrictions on infants' access to these phonological representations (cognitive, attentional, featural, temporal or otherwise) in mispronunciation detection or naming tasks.

Given the disparity between Stager and Werker's results and those of the studies mentioned above, researchers have grappled with understanding the factors that influence this phonological specificity, in addition to investigating the psychological reality of the two alternative explanations. These factors will be discussed in some detail in what follows, with the aim of outlining the consensus on the currently agreed explanation for the various findings in the literature.

3.1 Word Familiarity

One of the obvious differences between Stager and Werker's study and most of the studies mentioned above is that the words used in the former were introduced to infants in a laboratory setting, whilst the other studies tested infants with words they were already familiar with. Perhaps, restrictions on mispronunciation sensitivity are limited to newly learnt words and not to familiar words (Barton, 1976; Metsala, 1999; Swingley and Aslin, 2000). Contrary to this explanation, however, recent studies have found that infants at 14 months of age are sensitive to the mispronunciations of the vowels (Mani and Plunkett, 2008a) and consonants (Ballem and Plunkett, 2005) in newly learnt words (i.e. taught to infants just before the experiment), suggesting that even newly learnt words are encoded in sufficient phonological detail, and there are no cognitive limitations on infants' access to these phonological representations (see also Bailey and Plunkett, 2002 in this regard).

3.2 The Mispronounced Phonemic Segment

Whilst early work on this topic focused merely on whether or not infants displayed sensitivity to mispronunciations of words, more recent work has investigated the pattern of mispronunciation sensitivity to different kinds of phonemic segments. This work takes off from the theoretical standpoint advocated by Nespor et al. (2003) suggesting that consonants may be more integral to lexical recognition than vowels, while vowels may play a more crucial role in prosodic processing (given that prosodic emphasis is usually indicated on the vowels in words). Evidence for this perspective comes from work by Nazzi (2005), suggesting that 20 month olds can simultaneously learn two words that differ by merely a single consonant, while being unable

to learn two words that differ by a single vowel. Furthermore, while older children (30 months old) can simultaneously learn two words that differ by only a vowel, they still show a preference for consonantal rather than vocalic information in lexical acquisition.

Fikkert and colleagues present an alternative view to Nazzi's work using the notion of a Featurally Underspecified Lexicon (FUL) (Lahiri and Marslen-Wilson, 1991; van der Feest and Fikkert, 2005; Kager et al., 2007). The key characteristics of this model are that, at least early in infancy, specification of the place of articulation of a word is defined by the level of specification of the vowel. Second, lexical entries do not contain any specification of the feature [coronal]. According to FUL, infants fail to distinguish *bin-din* in the Stager and Werker (1997) study because of the lack of specification of coronal vowels like /i/. In contrast, infants readily discriminate *bon-don*, due to the conflict between the place of articulation of the labial/dorsal vowel and the coronal [d] in the heard word, *don* (Fikkert et al., 2005) – phonological specification of a word rests on the specification of the vowel in the word.

Furthermore, Curtin et al. (2009), find that infants as early as 15 months of age successfully learn two words that differ by a vowel in a habituation task, while failing to do so when the words differ by a consonant (Stager and Werker, 1997). Curtin et al. argue that this pattern of sensitivity is linked to the acoustic salience of the changes made, since infants in the task only succeeded in learning the acoustically most salient contrast (/i/-/ɪ/) but failed on acoustically less salient contrasts (/i/-/u/ and /ɪ/-/u/).

Results from word recognition tasks in infancy support Curtin et al.'s findings. Mani and Plunkett (2007, 2008a; Mani et al., 2008) report systematic comparisons of infants' sensitivity to single-feature consonant and vowel changes to familiar and novel words, and find no difference in infants' sensitivity to consonant or vowel mispronunciations of familiar words by as early as 12 months of age (Mani and Plunkett, 2010) and novel words by 14 months of age (the earliest age tested: Mani and Plunkett, 2008a). Like Curtin et al., Mani and colleagues find that this mispronunciation sensitivity is restricted by the acoustic salience of the mispronunciation – with acoustically less salient mispronunciations passing undetected relative to more salient mispronunciations (Mani et al., 2008; Mani and Plunkett, 2010; Mani and Plunkett, 2008b).

The difference between Nazzi's results and those reported by Curtin et al., and Mani and colleagues could, therefore, come down to the acoustic salience of the consonant and vowel changes tested in the different tasks. The concordance between results from Mani and colleagues, Curtin et al. (2009), and Narayan et al. (2010), therefore, suggests an important role for acoustic information in phonological development. However, phonological theory has long postulated a correlation between acoustic and phonemic features that might

have important implications for our understanding of the role of acoustic salience on phonological acquisition. In the next section, we examine research investigating the psychological reality of the phonological feature and the interaction between acoustic and phonemic features in lexical representation.

3.3 Features Contributing to Mispronunciation

The current conception of the phonological feature stems from the lineage of work by Jakobson et al. (1952; distinctive features) and Chomsky and Halle (1968; phonetic features), in themselves harking back to the classification of the articulatory characteristics of consonants by ancient Sanskrit grammarians (see Botma et al., this volume, for discussion). The phonological feature provides a tool for the description of the sounds of all natural languages, such that the presence of a positive value for a feature in the description of a sound lends itself to a corresponding articulation. While features serve eminently useful descriptive purposes, researchers assessing phonological development in infancy have recently begun to focus on the psychological reality of the phonological feature. In support of the phonological feature, White and Morgan (2008) report that 19 month olds display graded sensitivity to consonant mispronunciations of familiar words – that is, infants are more sensitive to 3-feature mispronunciations (place of articulation, manner of articulation, and voicing), than to 2-feature mispronunciations (place of articulation and voicing), which in turn are more salient than 1-feature mispronunciations (place of articulation) of word-initial consonants in familiar words. However, it is reasonable to assume that changes to two consonantal features may impact the acoustic characteristics of the mispronunciation more than changes to a single consonantal feature, given that features were often proposed as a descriptive tool for acoustic changes (Chomsky and Halle, 1968). While infants' graded sensitivity suggests understanding of the number of phonological features contributing to the mispronunciations, in the absence of acoustic analysis of the mispronunciations, it is hard to discriminate the contribution of acoustic and featural changes to the displayed mispronunciation sensitivity.

In light of the above argument, Mani and Plunkett (2008b) report a similar comparison of infants' sensitivity to 1-, 2- and 3-feature mispronunciations of vowels in familiar words. This work highlights two points of interest. First, infants show a graded sensitivity to vowel mispronunciations at 24 and not 19 months of age. One possible explanation for this developmental delay with vowels relative to consonants places the blame on the acoustic characteristics of vowels and consonants in speech. Since the production

of vowels tends to be acoustically more variable than consonants, infants may require greater exposure to vowel tokens in order to extract abstract phonemic features of their vowel categories, and consequently require more experience to display graded sensitivity to incremental vowel mispronunciations. Some support for the above explanation comes from Mani and Plunkett's (2008b) finding that infants' graded sensitivity to the different kinds of mispronunciations was better explained by the acoustic rather than the featural characteristics of the mispronunciations, thereby suggesting a crucial role for acoustic information in the detection of mispronunciations, and the representation and recognition of lexical items in infancy.

There is still hope for the phonological feature, however. For instance, phonologically, vowels tend to be described in terms of tongue height, tongue backness and the roundedness of the lips. In an assessment of the influence of different vocalic features on mispronunciation sensitivity, Mani et al. (2008) find that 18 month olds display sensitivity to changes to the height and backness, but not to changes to the roundedness of the vowel in familiar words. On the one hand, acoustic analysis suggested that roundedness mispronunciations were acoustically least salient. However, analysis of the vocalic repertoire of Southern British English infants suggested that roundedness is a largely redundant feature – given the high correlation between backness and roundedness. In the absence of evidence discriminating these two explanations, there is no *a priori* reason to suggest that the pattern of infant sensitivity reported is more or less feature-based rather than acoustic.

In addition, many researchers support the notion of FUL that predicts differences in infants and adults' sensitivity to phonemic changes based on the specific features changed (Lahiri and Marslen-Wilson, 1991; van der Feest and Fikkert, 2005, to appear; Kager et al., 2007). For instance, phonologists often characterize the feature [coronal] as unspecified, or the default place of articulation feature. Based on this unspecifiedness of [coronal] in the mental lexicon, Lahiri and colleagues (Lahiri and Reetz, 2002; Friedrich et al., 2008; Eulitz and Lahiri, 2004) report behavioural and neurophysiological evidence suggesting that infants and adults alike are not sensitive to mispronunciations of [coronal] place of articulation, despite displaying sensitivity to [coronal] mispronunciations of other places of articulation (i.e. participants detect the mispronunciation of *book* as *dook* ([labial] mispronounced as coronal), but not the mispronunciation of *dog* as *bog* ([coronal] mispronounced as [labial])). Since the differences in the acoustic characteristics of the mispronunciations are likely to be non-significant, Lahiri and colleagues attribute this asymmetry to the underspecification of some phonological features in the infant and adult lexicon, thereby, indirectly providing support for the psychological reality of the phonological feature.

3.4 Task-Related Effects

Recently, a number of studies have focused on task-specific effects in the infant literature. In contrast to Stager and Werker's finding that 14 month olds cannot simultaneously learn two similar-sounding word-object associations (i.e. *bih-dih*), Fennell et al. (2007) (see also Fennell and Waxman, in press) report that under specific circumstances, 14 month olds demonstrate robust learning of the *bih-dih* contrast. For instance, presenting the two novel words in referential phrases ('Look at the _____!') leads to infants successfully associating the two novel words to the correct novel objects. Fennell et al. (2007) argue that the phrases indicate to infants the referential nature of the novel words, and facilitates the association of the label and object tokens. Different tasks and experiment designs, therefore, place vastly different requirements on the infant cognitive system, and need not necessarily reflect differences in the underlying phonemic representations of sounds or words in infancy.

Nazzi's (2005) word-based categorization task requires infants to categorize three objects on the basis of the shared nomenclature of two of the objects (i.e. two novel objects are called *duk*, and one novel object is called *guk*). This requires an interaction between lexico-phonological information, working memory constraints and categorization abilities at a high level of processing. Inattention to vocalic information in Nazzi's task need not, therefore, reflect differences in the representations of vowels and consonants in the infant lexicon. Rather it may reflect differences in the constraints imposed by vocalic and consonantal information at this higher level of processing.

What is required is an analysis of the cognitive demands of the different tasks and correlation to the characteristics of vowels or consonants that cause their differential involvement in certain testing situations. Based on the importance of acoustic information in the studies outlined above, we might again implicate the variability in the acoustic characteristics of vowels in words relative to consonants. This variability might imply that it takes infants longer to develop abstract representations of the vowels in words, thereby leading to something akin to token-based discrimination of vocalic information in words. However, in the absence of evidence to this effect, we can only suggest that, at the very least, the underlying representation of vowels and consonants in the early lexicon is relatively robust, whilst susceptible to acoustic salience. Furthermore, the results reported above suggest that, at least in lexical recognition tasks, and in some word learning tasks, vowels and consonants equally constrain recognition even in early infancy.

3.5 Neighbourhood Density Effects

Much of the debate on the phonological specificity of early lexical representations stems from work by Charles-Luce and Luce (1990, 1995) suggesting that word recognition by young children does not access specification at the segmental level, and is based on a less than detailed analysis of speech content. We must note here that most of the results reported in the previous section do not address whether specification (or mispronunciation sensitivity) is at the segmental or word level – they seek, rather, to examine whether recognition is based on detailed analysis of the speech input.

Charles-Luce and Luce (1990, 1995) further specify that increase in the neighbourhood density of the words in the infant lexicon leads to an increase in the amount of acoustic detail involved in word recognition. For instance, they report that 81% of three-phoneme words in the average 5 year old's lexicon have zero neighbours. Based on these statistics, then, children would not need to pay attention to segmental level detail in order to differentiate words, and could survive on more global or holistic differences (see Storkel, 2002 for similar arguments). In contrast to this approach, however, Coady and Aslin (2003) suggest that entries in the early lexicon populate those areas in the adult lexicon with the highest neighbourhood density. Furthermore, neighbourhood density is actually higher in the younger lexicon, when expressed as a proportion of the lexicon (child or adult), suggesting that increasing neighbourhood density may not have an important role in increasing phonetic detail in the infant lexicon (if the infant lexicon is, indeed, phonologically underspecified).

4. Summary

The evidence reported here suggests an important role for different kinds of information in guiding early phonological development. Factors such as acoustic, articulatory and distributional information work to influence the fine-tuning of the infant phonological repertoire. Additionally, a number of models have been discussed that seek to explain the pattern of findings in the literature, notably, the Perceptual Assimilation Model/Articulatory Organ Hypothesis (PAM/AO) and the Neural Commitment Model/Native Language Magnets (NLM), with regard to infants' acquisition of the sounds of their native language. Between them, the two models explain most of the findings in the literature; however, it appears that PAM/AO may be more suited to characterizing the loss of sensitivity to non-native sounds, while NLM may be more appropriate to describing the development of sensitivity to native

language contrasts. Together, the models account for the influence of frequency, distributional, acoustic, and articulatory information on native language speech perception. One aspect of the findings that they do not cover though is the influence of cognitive development on phonological acquisition (but see Lalonde and Werker, 1995). Once again, this appears to influence native and non-native sensitivity in different ways (Conboy et al., 2008); raising the suggestion that explanation of these diverse findings may require different strategies. Sensitivity to non-native contrasts may depend on domain-general cognitive constraints, while sensitivity to native language contrasts may depend on more language-specific information, given infants' exposure to their native language.

Once the infants' native language repertoire is in place, there appear to be few restrictions on infants' access to this phonological inventory in representing the words they acquire in the second year of life. Any restrictions that are there appear to be modulated by either the acoustic or phonological salience of the phonemes tested, rather than by infants' familiarity with the words or the size of the infants' vocabulary. Indeed, there appears to be little support for the view that the neighbourhood density of early vocabularies influences the specificity with which lexical items are represented (especially given that infant neighbourhoods populate the denser sections of the adult lexicon: Coady and Aslin, 2003).

While infants' sensitivity to mispronunciations relies on the acoustic or featural characteristics of the words/sounds tested, differentiation between the roles of vowels and consonants in the early lexicon appear to affect higher levels of processing than the stages of lexical representation, for instance, the cognitive load imposed by categorization or mutual exclusivity constraints. Therefore, although vowels and consonants appear well specified in the infant lexicon, there may be differences in infant and adult access to these representations in different tasks, based on the cognitive limitations of the task. Yet again, these differences appear to operate at a domain-general level of processing, rather than linguistic processing, *per se*.

What is required, at this stage, is greater analysis of the separation between acoustic and phonological information in the infant lexicon, and investigation of infants' sensitivity to variation in acoustic and phonological input. While Mani and Plunkett (2008b) suggest that acoustic information may be more pertinent to explaining the pattern of infant sensitivity, one cannot ignore the fact that acoustic and phonological input tend to be highly correlated, thus making it difficult to assess their individual contributions to infant development.

In answer, then to the question, 'When does language begin?' – we conclude that at least as far as phonological development is concerned, language

experience begins to shape infant behaviour from day one (and before, as some studies suggest). While this may sound optimistic (indeed, we may not want the take-home message of this chapter to be that 5 month olds can read and perform complex mathematical calculations), we note that there are considerable limitations on the level of development caused by language exposure. Indeed, can such infant responding at 1-day-old be called linguistic, as opposed to a demonstration of their domain-general ability to encode and compare physical stimuli, which happen to be linguistic? This begs the question, of course, as to when language behaviour can be called linguistic.

5. Notes

1. While Oller argues for protophones, that is, speech-like characteristics in infant babbling, he is in favour of the position that this speech-like babbling is not language specific until later in life.
2. While the prosodic and segmental characteristics of infant babbling and adult speech may be similar, infants may not follow the temporal restrictions imposed by their native language, thus making babbling sound more different from adult language use.

12 Second Language Phonology

Heidi Altmann and Barış Kabak

Chapter Overview

1. Introduction	298
2. The L2 Acquisition of Segmental Structure	299
3. The L2 Acquisition of Suprasegmental Structure	305
4. Conclusion	317
5. Notes	318

1. Introduction

Acquisition of a second language (L2) phonological system projects multi-layered issues not only because it intersects with the complexities of human development, learning and the timeless laws of language, but also with what the learner brings into the acquisition process. The latter properties range from the learner's first language (L1) background, her age of acquisition, amount of exposure to the target language, and their undeniable interaction with an array of social and affective factors. Due to its interdisciplinary nature, research on L2 phonology has often approached empirical and theoretical constructs in phonology (e.g. phonological contrasts, the syllable etc.) through the angle of psycholinguistic principles governing their acquisition, retainment and loss by L2 learners. As such, the field of second language phonology has provided us with an invaluable tool to understand the nature and dynamics of phonological representations and processes, as well as universal principles governing phonological structures. In this chapter, respecting the traditional divide between segmental and suprasegmental phonology, we will deal with theoretical and psycholinguistic aspects of L2 phonology as they relate to the acquisition and processing of segmental (Section 2) and suprasegmental structure (Section 3). Regarding suprasegmentals, we focus on L2 syllable structure and phonotactic knowledge (Section 3.1) as well as L2 stress (Section 3.2.) with special attention

to recent L2 studies on stress perception and production, and structural factors in the acquisition of L2 stress.

2. The L2 Acquisition of Segmental Structure

Acquisition of phonology is not acquiring segmental inventories but *contrasts* that form the pillars of a phonological system. Research over the past couple of decades has convincingly shown that L2 phonological grammars are highly abstract, and should not be characterized as an (imperfect) imitation of L2 structures. Several factors may impede the realization of phonetic and phonological structures in the L2, which leads to what is referred to as ‘foreign’ or ‘non-native’ accent. Earlier attempts to disentangle such factors typically pursued the contrastive analysis methodology, whereby similarities and differences between the L1 and L2 were identified and predictions with respect to the rate and success of learning were posited on the basis of the projected strength of L1 transfer (e.g. Lado, 1957; Brière, 1966, 1968). In the 1970s, amid revolutionary ideas about L2 grammars as constituting natural languages and exhibiting unique properties traceable neither to the L1 nor L2 (e.g. Selinker, 1972), work on L2 phonology started to incorporate universal principles such as markedness (e.g. Eckman, 1977; see Major, 2001; Eckman, 2008 for a review) to understand the nature and dynamics of interlanguage phonological grammars.

Here, we approach the concept of ‘contrast’ essentially as a psycholinguistic construct, which hinges on at least two mechanisms (i) the realization (perception) of linguistically plausible (universal) oppositions¹ and (ii) the subsequent abstraction of the perceived contrast so as to encode it in some form of linguistic representation that closely approximates native speakers’ representations. While (i) is a prerequisite for (ii), they are not necessarily isomorphic, as, for example, in the case of accurate production of segmental and prosodic contrasts that are not part of the learner’s L1, or in cases where learners are able to perceive the contrast but unable to encode it.² One can thus approach foreign accent as a multitude of factors that lead to realizations of imperfect (or incomplete) representations due to an inability to accurately (native-like) perceive L2 contrasts and/or failure to lexicalize them. Researchers have focused on both the perception and production of L2 structures from a theoretical, psycho-acoustic, psycholinguistic as well as functional point of view.

In this chapter, we frame the empirical discussion of L2 phonology largely in terms of research on L2 perception and production, noting how the study of each contributes to our understanding of L2 learners’ phonological and phonetic competence. Due to space limitations, we leave aside several learner-related factors (e.g. age of acquisition, amount of exposure, affective factors etc.) which undeniably influence the L2 acquisition of sound structure.³ We also

exclude the discussion of the phonetic aspects of second language perception and production, cross-linguistic variability in attending different phonetic cues in the detection of phonological contrasts, as well as the issue of intelligibility in L2 speech.⁴ The primary research questions that are of interest to us here are: Which structural factors hinder or enable the accurate perception of L2 phonological contrasts? How do L2 phonological representations differ from those of native speakers?

2.1 Similarity-Difference, Markedness and the Role of Perception

In comparison to other domains of grammar, structural similarity and difference between L1 and L2 sound systems have been, and continue to be, at the heart of theorizing in L2 acquisition of segmental contrasts. Eckman (1977), in his *Markedness Differential Hypothesis*, however, asserts that the learning of L2 contrasts can be predicted to be difficult if they are not just different from the L1 but also relatively more 'marked' than the L1 (implied here is typological markedness as, for instance, operationalized by implicational hierarchies). For example, on the basis of Cantonese learners' production of English voicing contrasts, Eckman (1981) showed that the interlanguage of Cantonese speakers exhibits the rule of terminal devoicing, whereby the voicing contrast in all word-final obstruents is neutralized (e.g. [pɪk] for 'pig'), although Cantonese has no voiced consonants and consequently no L1 forms to motivate the phonological rule in question. Furthermore, the target language (English) also lacks this rule. As such, the interlanguage of the Cantonese speakers involves a phonological process that has no role in either the L1 or L2. What seems to determine such an 'interlanguage' rule then is attributed to the notion of markedness: Voicing contrasts word-finally are more marked than those word-medially and/or word-initially. The interplay of markedness with language-specific structural properties has recently attracted a remarkable amount of interest within theoretical work on L2 phonology, primarily influenced by the ideas developed within the framework of Optimality Theory (e.g. Broselow, 1999; Broselow, Chen and Wang, 1998), which we will address below in more detail.

Structural differences, however, are relative, and can stem from a number of sources. For example, Flege's *Speech Learning Model* (Flege, 1987, 1990, 1995) postulates testable hypotheses which, among other things, assert that the likelihood of perceiving the difference between an L1 and L2 sound increases with decreasing similarity between the respective L1 and L2 sound. Accordingly, L2 learners are predicted to establish a new L2 sound category if they perceive all or at least some differences between the L2 sound in question and its closest L1 counterpart. Flege's (1987) study on the acquisition of French by English learners confirms this prediction: English learners of French were shown to produce

the French /y/ (high, front, rounded vowel), which is absent in English, more authentically than the French /u/ (high, back, rounded vowel), which is very similar (but not identical) to its English counterpart (Flege, 1987). Flege's notion of 'equivalence classification', a sound substitution strategy where perceptually linked sounds are assimilated into a single category, is argued to block the formation of new L2 categories and consequently to hinder the acquisition of L2 contrasts.

Eckman, Elreyes and Iverson (2003) suggest at least three different situations where L1 and L2 may employ different phonemic contrasts. This is summarized in Table 12.1.

Table 12.1 Possible differences between L1 and L2 sound contrasts (adapted from Eckman, Elreyes and Iverson, 2003)

	L2	L1	Example
1	Sounds A and B are in contrastive distribution	contains neither A nor B	Korean has neither /f/ nor /v/, which are both contrastive in English.
2		contains A, but lacks B	Japanese has /p/ but lacks /f/, which are both contrastive in English.
3		contains both A and B as allophones of the same phoneme	The phonemes /d/ and /ð/ are in allophonic distribution in Spanish while they are in contrastive distribution in English.

Despite the fact that, in comparison to Situations 1 and 2, Situation 3 illustrates a difference where the L2 appears to be the most similar to the L1, there is an abundance of evidence that suggests that Situation 3 seems to cause more difficulty than the other two situations (e.g. Lado, 1957; Stockwell and Bowen, 1965; Hammerly, 1982; Hardy, 1993; Flege, 1987, 1990; Major and Kim, 1999; Eckman, Elreyes and Iverson, 2003). Presence or absence of an L2 sound in the L1 system cannot then adequately explain why speakers of different L1s with comparable sound inventories face different levels of difficulty with the same novel contrast. In this respect, we need to look for an explanation rooted in the way learners handle the speech signal. On the one hand, there is a large body of evidence that strongly suggests that infants as young as 1 month old are able to acoustically distinguish both ambient and target language contrasts (see Mehler, 1985 for a review). On the other hand, this remarkable ability decreases rapidly in infancy with exposure to the native language (see Werker and Polka, 1993; Mani, this volume, for a review). It is therefore reasonable to assume that an already established native language phonological system can continue to exert its influence on speech perception in adulthood. Such an influence can

become especially evident when adults encounter a new phonological system, as in the case of L2 acquisition. Hence, the interrelation between speech perception and phonological acquisition can provide a more satisfactory explanation for why we observe a great deal of variation among L2 speakers' engagement with different types of contrasts, ranging from chance-level to near-native-like performance (Best et al., 1988; Polka, 1991; Pruitt, 1995; see Brown, 2000 for a model that links infant and adult speech perception to account for the influence of L1 grammar in L2 acquisition of phonology). Furthermore, 'phonetic' factors such as phonetic exposure to the members of a contrast in non-contrastive contexts (e.g. Burnham, 1986; Harnsberger, 2000; Ingram and Park, 1998) as well as the psycho-acoustic saliency of the contrast (e.g. Burnham, 1986; Kabak and Maniwa, 2007) may also influence the accurate perception of non-native contrasts. In lieu of this, a number of models have been proposed in the literature that attribute the degree of perceptual difficulty to how well the members of the non-native segments in contrast are heard as instances of existing L1 categories. Naturally, the quantification of 'phonetic similarity' varies from one model to another. For instance, *The Perceptual Assimilation Model* (PAM) (e.g. Best, 1993; Best, 1995; Best et al., 1988; Best et al., 2001) asserts that if an L2 sound is similar to an L1 sound, it is assimilated to the corresponding L1 category. However, if an L2 sound does not have any detectable similarities to any sound in the L1, the sound is uncategorizable (i.e. it falls within native phonological space but in between existing categories). There are times when an L2 sound might be too distant from any L1 sound, as would be in the case for Zulu clicks for native English speakers, such that it can be perceived as a non-linguistic sound. PAM proposes 6 different degrees of perceptual differentiation for L2 contrasts and predicts different discrimination levels for each:

1. *Two-category assimilation*: Each L2 phoneme is assimilated to a different L1 category. Discrimination is expected to be excellent.
2. *Category-goodness difference assimilation*: Both L2 phonemes are assimilated to the same L1 sound but one is heard as a better instance of the L1 category than the other. Discrimination is expected to be moderate to very good.
3. *Single-category assimilation*: Both L2 phonemes are assimilated to the same L1 category and are heard as similarly good or poor instances of the L1 segment. Discrimination is expected to be poor.
4. *Both uncategorizable assimilation*: Both L2 phonemes are heard as possible sounds within listeners' native phonological space but cannot be assimilated to any L1 phoneme. Discrimination is expected to range from poor to very good, depending on their proximity to each other and to L1 categories within the L1 phonological space.

5. *Uncategorized versus categorized assimilation*: One of the L2 phonemes is assimilated to an L1 category and the other is not. Discrimination is expected to be very good.
6. *Nonassimilation*: Neither of the L2 phonemes is recognized as a sound that could be produced within listeners' L1 phonological space. Discrimination is expected to be good to very good.

In a similar vein, the *Native Language Magnet Model* (NLM) (e.g. Kuhl, 1991; Iverson and Kuhl, 2000) equates L2 speech perception to a distorted lens through which language is filtered. The major claim of the model is that a 'perceptual magnet effect' is responsible for the changes in phonetic perception, which 'warps' the acoustic space underlying phonetic perception such that not only the way we perceive speech but also the way we produce it get affected. As such the NLM provides a principled explanation as to why phonetic similarity, as discussed above, might be problematic. More specifically, NLM introduces the notion of a 'prototype', roughly a very good instance of a particular sound category. In the case of L2 learning, L1 prototypes may act as a perceptual magnet, attracting less prototypical instances of the same category – that is, similar L2 sounds – and consequently reducing the perceptual distance between the L2 sounds and the prototypical L1 sounds.

In summary, the success at which L2 phonemic contrasts are acquired seems to largely depend not only on the straightforward correspondence between L1 and L2 sound inventories, but also on the degree of perceived similarity and difference on the part of the L2 learner.

2.2 Can Novel L2 Contrasts Be Ultimately Acquired? Lexical Encoding and the Nature of L2 Lexical Representations

Clearly, L1 acquisition and L2 acquisition are qualitatively and quantitatively different: adult L2 learners rarely achieve the native competence that children do when they master their L1. The consistent lack of convergence on the target language by L2 learners may be attributed to the fact that Universal Grammar (UG), which presumably gives a 'free-ride' to the learner in the acquisition process, does not operate in L2 acquisition.⁵ However, other factors with regard to input characteristics, learning styles and, most conspicuously, the fact that the L2 learner tackles the acquisition process already knowing a language provide more plausible explanations for the L1 versus L2 acquisition dichotomy, rather than just the inaccessibility of UG. In L2 phonology literature, these latter factors, more specifically, the extent of L1 transfer and its relation to L1- and L2-external factors (e.g. markedness) have defined the research agenda

(see Ioup and Weinberger, 1987): Can new phonological structures that are absent in the L1 be successfully acquired after the L1 system is fully established? How exactly do L1 transfer and UG operate in the acquisition of phonology? Does the absence of an L2 contrast in the L1 always imply an inability to learn it?

In this respect, Brown (1998) argued that the inability to acquire new phonological representations is not due to the inaccessibility of UG, but rather to 'insufficient intake' to the language acquisition device. Insufficient intake occurs due to L1 transfer, which heavily influences speech perception. Brown compared Japanese and Chinese learners of English in their acquisition of the /l/-r/ contrast. Crucial to this study is the fact that both L1s lack the contrast in question. While the Japanese learners performed approximately at chance level, Chinese learners perceived the distinction as successfully as native speakers of English. The asymmetry between the two non-native populations for Brown was related to the language-specific employment of the phonological feature relevant for the L2 contrast in question. More specifically, she argued that the English liquid contrast crucially hinges on the feature [coronal], which is not present in Japanese but is in Chinese.⁶ The latter is argued to actively employ this feature to distinguish other sounds in its system, such as the contrast between the alveolar (/s/) and the retroflex (/ʂ/) fricative. As such, while the Chinese learners were able to build new phonological representations by recycling the already existing and active [coronal], the Japanese learners could not have access to this recycling utility since the feature has no contrastive function in the L1. While other work provided evidence in support of this view (e.g. Matthews, 1997; Atkey, 2000; see Archibald, 2004 for a summary), LaCharité and Prévost (1999) argued that not all inactive L1 phonological features pose learnability problems for those segments that they distinguish in the L2. Accordingly, the researchers made a distinction between articulator features (i.e. organizational nodes, which are not allowed to occur terminally, such as [coronal], [dorsal] or [pharyngeal]) as opposed to those features that occupy the terminal nodes (i.e. those that are dependent on other features such as [distributed], [anterior] or [back]). In their study, Canadian French speakers did better with the English /θ/, which is distinguished by the terminal feature [distributed], than /h/, which is represented with the articulatory feature [pharyngeal]. In the same vein, Curtin, Goad and Pater (1998) showed that English speakers learning Thai can indeed learn Thai three-way stop voicing contrasts, which depend, among others, on the feature [aspiration], although the same feature is inactive in English (see, however, Pater, 2003, who was unable to replicate these findings). More cross-linguistic research is however necessary to further test the validity of feature-based accounts on the interaction between phonological features and phonetic cues that are necessary to trigger the acquisition of novel contrasts. It should be noted that there is growing experimental evidence that suggests that learners often improve

their ability to perceive non-native contrasts, including the notorious English liquid contrasts by Japanese speakers (e.g. Strange and Dittmann, 1984; Bradlow et al., 1997, 1999), which questions the generalizability of such accounts.

However, little is known about the precise lexical encoding of problematic L2 contrasts by L2 speakers. Are words that minimally differ with respect to 'difficult-to-perceive' contrasts (e.g. *lack* versus *rack* by Japanese speakers) encoded as homophonous in the L2 learners' lexicon? Pallier et al. (2001) observed repetition priming⁷ for Catalan minimal pairs differing by contrasts that only exist in Catalan (e.g. /o/ versus /ɔ/) by Spanish-dominant Spanish-Catalan bilinguals but not by Catalan-dominant Spanish-Catalan bilinguals, suggesting that Spanish-dominant bilinguals accessed both members of the minimal pairs upon hearing either word. Recently, Hayes-Harb and Masuda (2008) showed that learners, with increased exposure to L2, can store distinct lexical representations built on novel contrasts (e.g. the geminate versus singleton contrast in Japanese such as /t/ versus /tt/). This success, however, appears to be specific to listening tasks as opposed to production tasks whereby the same learners did not achieve the same level of performance as a consequence of neutralizing the contrast in question (see also Mah and Archibald, 2003 for similar observations). These results can be taken to suggest that when learners encounter novel phonemes in the auditory input, they may be able to successfully distinguish them solely on the basis of the fact that they know a difference, though unfamiliar, exists (e.g. [tt] as being the unfamiliar variant of [t]). Inaccurate production of the same contrast, hence, is not surprising since this sort of lexical encoding does not accurately include knowledge as to how the new contrast is phonetically implemented (see also Cutler, Weber and Otake, 2006). The production-perception asymmetry in L2 is also observed in the acquisition of L2 stress, as we will discuss in Section 3.2.

3. The L2 Acquisition of Suprasegmental Structure

Phonological knowledge not only involves the knowledge of segmental contrasts, but also their distribution and further organization into prosodic units such as syllables and feet, which form the building blocks of the phonological representations of words, and eventually into higher prosodic structures at phrase and utterance level. Just as L1 segmental phonology can exert its influence on the acquisition of L2 contrasts, L1-specific suprasegmental phenomena also seem to pervasively dictate the way L2 sound strings are perceived and pronounced. When such higher-level phonological structures emerge in language-contact situations and filter language processing, an even more compelling case for the complexity and abstractness of phonological knowledge is made. Below, we will focus only on syllable structure and word stress, as these

structures constitute two of the most widely studied suprasegmental phenomena in L2 albeit research on L2 prosody is still at its infancy. The reader is referred to Chun (2002) and Trouvain and Gut (2007) on the L2 acquisition of intonational phonology; Wang, Jongman and Sereno (2006) on the L2 acquisition of tonal phenomena, and Trofimovich and Baker (2006) on the fluency characteristics of L2 speech.

3.1 L2 Syllable Structure and Phonotactic Knowledge

Languages differ not only with respect to their segmental inventory but also in the way they combine the sounds that exist in their phonological system. For example, while both Turkish and English employ /h/ contrastively, this sound can never be realized in the coda position in English while Turkish allows the occurrence of the same sound in both the coda and onset (e.g. *tah.ta* 'wood' versus *še.hir* 'city' – where <> indicates a syllable boundary). Syllable complexity is yet another difference across the languages of the world. For example, while Japanese and Korean do not allow complex onsets and codas, the same is banned only from the onset position in Turkish while the coda position tolerates some type of consonant clusters (e.g. *Türk* 'Turk', *çarp* 'multiply').

In the SLA (Second Language Acquisition) literature on syllable complexity, it has been suggested that both the L1 and the universal segmental sequencing properties (i.e. phonotactics) can influence the L2 learners' performance. The rate and difficulty at which such structures are acquired have typically been predicted by a comparison of these two properties. For instance, Broselow and Finer's (1991) *Minimal Sonority Distance* (MSD) parameter, based on Selkirk (1982), characterizes variation in cluster sequencing based on the degree of sonority difference between adjacent segments for each language.⁸ Roughly, a language that does not permit consonant clusters would have an MSD setting of 5, the largest allowable difference between neighbouring segments being 4. A language allowing [Obstruent+Glide/Liquid] sequences has the MSD setting 3 since the sonority difference between the members of the cluster is at least 3. Broselow and Finer proposed that L2 consonantal sequences with a lower MSD setting than the L2 learners' L1 setting will be difficult to acquire. As grammars arguably prefer larger differences with respect to sonority in complex clusters, it is also expected that given two L2 consonantal sequences, the one with the lower MSD will be the more difficult to acquire. As such, this account explains developmental effects in syllable structure through L1-specific properties as well as universal generalizations on sonority sequencing. Indeed, Broselow and Finer's (1991) experimental data from Japanese and Korean speakers of English indicated an upward trend in error rate as the sonority difference between consonant clusters decreased (e.g. [C]) clusters had a lower error rate than [Cr] ones

because the former has a higher MSD score than the latter). A later study by Hancin-Bhatt and Bhatt (1997) examined those cases where the L2 learner can unambiguously transfer complex onsets, as in the case of Spanish learners whose L1 MSD setting is 2, and extended the MSD's coverage by looking at the SLA of syllable codas, firmly supporting the predictions of the MSD model.

Other researchers such as Eckman and Iverson (1993), however, argued that the MSD parameter of Broselow and Finer (1991) makes incorrect predictions on a number of grounds. For instance, since Broselow and Finer argue that Japanese and Korean have [Cj] clusters, including [nasal+glide] sequences (e.g. Korean: [mjən] 'cotton', Japanese: [mjo:] 'strange')⁹, then an MSD setting of 2 must be assumed. As such, these languages should be able to allow a larger array of consonant clusters, but they do not. Adopting Clements' (1990) *Sequential Markedness Principle*, Eckman and Iverson instead argued that relative markedness relationships between sounds are more crucial. For example, voiced consonants are more marked than voiceless ones, and fricatives are more marked than stops. Since marked structures are predicted to be more difficult to acquire than unmarked ones (Eckman, 1977), consonant clusters with voiced segments or those with fricatives should pose more difficulty than those with voiceless segments or those with stops, respectively. Since consonant clusters with profiles containing severe sonority slopes are assumed to be more marked, [Cj] clusters are considered to be more marked than [Cr] clusters, explaining why languages like German and Spanish have the latter but not the former. The opposite would be predicted in Broselow and Finer's model since the MSD setting is higher for [Cj] clusters than [Cr] ones. While researchers differ in their quantification of syllable complexity, sonority sequencing, which stems from typological observations, seems to play an important role in the development of L2 syllable complexity.

3.1.1 *Interlanguage Rules Unmotivated by Neither L2 nor L1*

How do universal patterns emerge in interlanguage phonological grammars when the learner has no evidence for them in her L1? Recently, this question drew the attention of phonologists primarily working on Chinese learners of English and loanword adaptations in the 1990s (e.g. Wang, 1995; Silverman, 1992).¹⁰ Essentially, there is a tendency to insert a final vowel when the L2 form or the loan is monosyllabic and ends in a consonant, but to delete the same word-final consonant in bisyllabic words. Wang (1995) argued that the choice of this strategy by Mandarin speakers is related to a preference for bisyllabic forms. However, the tendency for bisyllabic forms in the interlanguage grammar cannot come from the L1 since languages like Mandarin and Cantonese allow monosyllabic words. The presence of such processes in L2 grammars that are not motivated by surface alternations in the L1 or L2 provided an instructive case to demonstrate the power of Optimality Theory (OT).¹¹ For instance, Broselow,

Chen and Wang (1998) argued that Universal Grammar provides speakers with a set of markedness constraints, which exert their forces by modifying underlying representations (e.g. via epenthesis, deletion, devoicing etc.) to yield less-marked optimal forms. When a Mandarin speaker is confronted with English forms ending in obstruents (e.g. *bed*), the learner, presumably at an earlier stage of development, will not be able to produce the input faithfully, hence resort to a repair strategy. The repair option is either the deletion of the final consonant or epenthesis, both of which incur an equal amount of faithfulness violations. The choice between the two repair strategies is determined by a so-called latent constraint that favours bisyllabicity (owing to other cross-linguistically dominant effects such as foot binarity). Since OT assumes that universal principles underlie all grammars, and they are part of the grammar of each language, the emergence of the unmarked in L2 can be straightforwardly accounted for, unlike rule-based accounts where phonological grammars are assumed to be a set of learned rules that convert underlying representations into surface representations. Upon encountering conflicting input in language-contact situations, such OT constraints become visible (see Broselow, 2004b, who applies a probabilistic constraint ranking approach in the spirit of Boersma and Hayes (2001) to account for variation in L2 productions).

3.1.2 The Role of L1 Phonotactic Knowledge

When speakers encounter sequences that are illicit in their L1, they adopt different repair strategies to fit them into their L1 phonotactics, both in the domain of perception and production. While repair processes occurring in the latter domain are relatively well known, owing to the abundance of (sometimes anecdotal) examples coming from loan word adaptations in Japanese and Korean, where we see the emergence of epenthetic vowels in otherwise illicit consonantal sequences (e.g. Korean: [k^hu.ri.su.ma.su] 'Christmas', Japanese: [ma.ku.do.na.ru.do] 'Mac Donald's'), research on the influence of phonotactic knowledge on perception has mostly focused on perceptual response biases favouring the interpretation of ambiguous sound combinations as legal sequences (e.g. Massaro and Cohen, 1983; Hallé et al., 1998; Pitt, 1998; Moreton, 2002). Factors contributing to phonotactic knowledge, such as frequency and lexical effects, have also been studied (e.g. Luce and Pisoni, 1998; McClelland and Elman, 1986; Pitt and McQueen, 1998; Vitevitch and Luce, 1998), altogether demonstrating that listeners' knowledge about the legality and the probability of phonotactic patterns matter in the processing of spoken stimuli.

Since phonotactic regularities of the ambient language have been shown to also influence infants' perceptual abilities (e.g. Jusczyk et al., 1993, 1994), an obvious question arises as to whether they can also play a role in L2 speech perception. Dupoux and colleagues asked whether the observed epenthesis effects in Japanese can also hold for non-native speech perception

(Dupoux et al., 1999, 2001; Dehaene-Lambertz et al., 2000). For instance, comparing Japanese listeners with French listeners in their perception of consonant clusters, Dupoux et al. (1999) employed a series of six items created from naturally produced nonce words (e.g. [abuno], [akumo], [ebuzo], [egudo] etc.) by gradually reducing the duration of the vowel [u] from its full state down to zero milliseconds. When participants were asked to respond whether the items they heard contained the sound [u], Japanese listeners, unlike French listeners, overwhelmingly judged that the vowel was present at all levels of vowel length. This was even the case about 70% of the time even when there was no vowel in the stimuli. Such results convincingly show that the influence of L1 phonotactics can be so robust that listeners 'invent' illusory vowels to accommodate illicit sequences of segments in their L1.

The aforementioned studies on Japanese leave the question open as to whether perceptual epenthesis stems from the fact that Japanese speakers never hear consonant clusters in their language or from the fact that Japanese syllable structure is predominantly CV. Kabak (2003) and Kabak and Idsardi (2007) replicated and extended Dupoux et al. results to Korean L2 speakers of English residing in the United States. Korean provides a more instructive array of consonantal contact restrictions. Unlike Japanese, which licences very few coda consonants and thus displays a paucity of coda-onset clusters (Itô, 1986, 1989), Korean provides potential contexts for heterosyllabic consonantal contacts since consonants belonging to different morphemes can come into contact at morpheme junctures. For instance, while [k] and [t] can occur in the coda position, strident consonants such as /c/ cannot be licensed in the same position. However, due to the processes of nasalization and lateralization, both [k] and [l] obligatorily undergo assimilation. Hence, while [k.t] and [l.t] sequences are perfectly acceptable (since the nasalization and lateralization rules of Korean do not apply to them), it is not possible to find sequences such as [k.m] and [l.n] on the surface (unless separated by a phrase boundary). This is because, due to the aforementioned assimilatory processes, such sequences must surface as [ŋ.m] and [l.l], respectively. As such, Korean provides a compelling test case where language-specific co-occurrence restrictions (e.g. constraints such as *[k.m], *[l.n]) can easily be separated from prohibitions entirely based on syllable structure restrictions (e.g. [c] cannot occur in the coda position). The results from an AX discrimination task¹² reported in Kabak (2003) and Kabak and Idsardi (2007) suggest that language-specific co-occurrence restrictions do not explain the (lack of) perceptual epenthesis phenomena in words containing sequentially illicit consonantal sequences. That is, words with illicit sequences such as **pakma* and **palna* were successfully discriminated from their epenthetic counterparts by Korean speakers of English. Rather, perceptual epenthesis was evoked when the members of the illicit sequences incurred a syllable structure violation, hence words with illicit sequences such as **pacma*, **pacta* could not be

discriminated from their epenthetic counterparts. The interpretation of these L2 results provides compelling evidence against phonological theories and analyses that employ syllable-independent and linear statements to explain the consonantal distributions (e.g. Dziubalska-Kořaczyk, 1994; Steriade, 1999a, 2001a; Blevins, 2002). Furthermore, sequences such as **pakma* and **palna* are not perceptually confused with *paŋma* and *palla*, which would be the expected surface realizations of such sequences in Korean production, demonstrating that production is not necessarily isomorphic to perception, and Korean listeners' L2 perception does not mirror their production habits in the L1.

In the spirit of parsers that are traditionally built for syntax (e.g. Phillips, 1996), L2 data on perceptual repair strategies have motivated the development of phonological parsers for L2 phonological systems (e.g. Archibald, 2004). For example, Kabak (2003) proposes a syllabically conditioned perceptual algorithm for L2, which incrementally builds the syllabic representations of input forms with reference to universal tendencies pertaining to syllable structure (e.g. sonority) as well as language-specific restrictions on the syllable (e.g. coda restrictions, allowance for branching codas and onsets etc.). The perceptual algorithm detects [Onset] and [Coda] information in the acoustic flow using 'onset' and 'coda' detectors. Roughly, these detectors search for the acoustic correlates of phonological features that are associated with particular prosodic positions in the language. As the algorithm builds the syllabic representation of the speech stream, it assigns [Onset]-interpreted features to onset positions, and [Coda]-interpreted ones to coda positions. Results from Korean L2 speakers of English suggest that when the perceptual algorithm encounters an L2 string such as *Ri[cm]ond* 'Richmond', it is forced to place the strident in an onset position. Since onsets are non-branching in Korean, [cm.] in the onset would be illicit. This motivates an epenthetic vowel in the illicit cluster in order to satisfy the basic syllable structure condition.

Yet another equally crucial outcome from these perceptual epenthesis studies is that Korean L2 speakers' perception of epenthetic vowels cannot be explained by the frequency of the sequences that are involved since certain consonant clusters ([k.m] and [l.n]) were successfully discriminated by Korean listeners although they have zero probability of occurrence in Korean. On a par with this result, Davidson (2006) also shows no correlation between the lexical frequency statistics and the accuracy with which monolingual English speakers produce pseudo-Czech word-initial consonant clusters.

3.2 L2 Stress

So far, we have shown that L2 learners may modify segmental as well as syllabic structure in perception or production. The same also holds true for the

placement of L2 word stress. Research in this field can still be considered to be in its beginning stages, and it was not until very recently that the investigation of L2 stress phenomena has been vigorously approached with large-scale studies on learners with different L1 backgrounds. It should be noted, however, that rather little is known so far about cross-linguistic effects that can be generalized to L2 acquisition in general in this area, that is, beyond the L2 acquisition of English stress. The focus in the 1990s was on L2 production, where learners' modifications could be easily discerned in incorrectly stressed words (Archibald, 1993, 1994, 1998; Pater, 1993, 1997). This, in turn, triggered interest in the perceptual side of L2 stress and led to the postulation of the 'stress deafness' hierarchy (Dupoux and Pallier, 1997, 2001; Peperkamp et al., 1999; Peperkamp and Dupoux, 2002) and motivated L2-specific in-depth cross-linguistic investigations (Altmann, 2006; Kijak, 2009) that attempted to directly relate the ability to perceive stress to different properties of the speakers' L1 prosodic system. In the following, we focus on the interaction between L1 prosodic parameters in stress production and perception, as well as on the influence of target-language-internal structural factors on the acquisition of L2 stress.

3.2.1 *The Production of L2 Stress: The Role of L1*

The initial focus of research in the area of L2 stress was to investigate if specific L1 parameter settings for stress can be *reset* in the course of L2 acquisition. The data of early studies indicated (or could be reanalysed in this direction) that L1 stress parameters seem to influence the way the learner places stress on L2 words. Consequently, the learner only displays target-like L2 stress placement in words where it coincides with L1. This was found, for example, in the context of L2 English for Jordanian Arabic (Anani, 1989) and Egyptian Arabic (Youssef and Mazurkewich, 1998) learners, as well as for Hungarian and Polish (Archibald, 1993) learners. However, since real existing English words were used as stimuli in these studies, it is unclear what kind of knowledge was actually tapped in these production experiments. Hence, these studies were not able to precisely answer what kind of knowledge learners have of the general L2 stress pattern. In addition, learners also displayed incorrect production patterns that could not be attributed to L1 transfer (labelled *parameter missetting*), or even followed no systematic pattern (Archibald, 1997).

More insight could be gained with the investigation of nonce words. For example, Korean learners were found to have much higher correctness scores for the production of known nouns and verbs in English than for novel ones (Guion, 2005). One of the first systematic studies that used nonce words (and native L1 controls) was Pater (1997) involving Canadian French learners of English. Stress was produced mainly at the left word edge by these learners, which corresponds neither to the L1 (French has regular stress on the right word edge) nor the L2 pattern. This indicated (i) learners' awareness that the

L2 stress pattern is different from the L1 pattern, and (ii) active application of some (although non-target like) L2 strategy rather than stress placement on the basis of memorized lexical items.

To expand the scope beyond specific single L2 populations, Altmann (2006) tested the L2 production of eight typologically different L1 groups for English nonce words. This study yielded two clearly distinct groups of L1s regarding strategies for L2 stress placement. The first group consisted of speakers of L1s without word-level stress (Chinese, Japanese, Korean). They performed most unlike the native English controls with no apparent strategy or pattern (e.g. consistent stress on the final syllable). The second group was comprised of speakers of L1s with phonologically regular stress (French, Turkish, Arabic), who surprisingly yielded most target-like patterns. This indicates that the existence of word-level stress in the L1 is a factor that may help learners with the application of target-like stress patterns in production. More specifically, the second group produced stressed syllables, which involved a distinct combination of pitch, loudness, and duration on the vocalic element, in a position and in a way that accurately complies with the target norms. It could, however, not be determined if non-target like productions of the first group were due to a mere failure to phonetically realize stress (i.e. accurately producing the expected L2-specific combination of the three basic acoustic correlates of stress) on the syllable that should have been stressed, or a general (phonological) failure to correctly decide which syllable should have been stressed.

Since the vast majority of statements about L2 stress acquisition (and this is also true for the field of L2 Phonology as a whole) are based on English as target language, Kijak (2009) investigated the acquisition of Polish as L2. Polish is a language with regular stress assignment, where stress is fixed on the penultimate syllable. In contrast, the English stress system is arguably not phonologically regular since stress is contrastive (e.g. *récord* versus *recórd*), so the addition of a target language with non-contrastive stress could provide more insight into the process of L2 stress production in general rather than only into that of English. In Kijak's results, a general pattern emerged that somewhat differs from that found by Altmann (2006): speakers of L1s that allow stress on the penultimate syllable and have contrastive stress (German, Spanish/Italian, Russian, English) showed highly target-like productions, and outperformed speakers of L1s that disallow stress on this syllable (i.e. those that have no word-level stress (Chinese) or fixed stress on a different syllable (French and Czech)). This latter group of L1s further split into French speakers, on the one hand, and Chinese and Czech speakers, on the other hand, with French outperforming the other two. However, level of proficiency also turned out to be a factor – higher-level learners produced more penultimate stresses than lower level learners for each L1, which indicates that accurate production of stress can be learned.

In all of the larger scale studies described above, both the perception and the production of stress were tested, which is crucial for solid argumentation. Trivially, only for learners with good L2 stress perception, good L2 stress production might be expected. We will see below, however, that learners who fare well in L2 production sometimes perform worse in L2 stress perception, while ones who have problems in production are able to perceive L2 stress at ceiling level. Thus, the relationship between perception and production in L2 is not a straightforward one, as already pointed out for the L2 acquisition of segmental structure in Section 2.

3.2.2 *The Perception of L2 Stress: The Role of L1*

In comparison to the production side, much more is known about the perception of L2 stress. As a starting point, in investigations using real words, (at least) a tendency could be observed that certain learners of English have a harder time identifying the location of word stress than native speakers (e.g. for Hungarians: Archibald, 1993; for Egyptians: Youssef and Mazurkewich, 1998). Given the prosodic properties of the learners' L1 in these cases (i.e. regular stress assignment), such findings provide support for an effect that has been labelled *stress deafness*. In a series of general psycholinguistic (i.e. not specifically L2-related) experiments, Dupoux and colleagues (Dupoux and Pallier, 1997, 2001; Peperkamp et al., 1999; Peperkamp and Dupoux, 2002) tested the perceptual abilities of speakers of languages that have regular (i.e. non-contrastive) stress for segmental and stress contrasts in trisyllabic words with open syllables. Based on their findings, they devised a *stress deafness hierarchy*, along which a significant but gradual decline of perceptibility of stress contrasts (e.g. *fidape-fidápe*) compared to phonemic segmental contrasts (e.g. *fidape-lidape*) can be modelled. This is argued to be in accordance with increasing transparency (surface-observability) of a language's stress properties for children in the course of their first language acquisition. If, for instance, word stress is fixed in initial or final position and thus will necessarily coincide with utterance edges, this can be inferred by children in the earliest stages of L1 acquisition and stress will not be encoded in their L1 grammar. If, however, stress is not always or never utterance-final (and thus less surface-observable), this requires further segmentation of utterances and thus leads to a later setting of the stress parameter in L1 acquisition (Peperkamp, 2004). In support of this claim, French and Finnish speakers, in whose L1 stress always falls at a word edge (and thus also at utterance edges), were found to show the least success in perceiving stress differences, while Spanish speakers, whose L1 employs contrastive stress (i.e. stress is not phonologically fixed on one given position within the word), were most successful. Polish and Hungarian, where stress is regular but not always at an utterance edge (e.g. due to regular penultimate

stress in Polish, or due to the existence of unstressed phrase-initial function words in Hungarian, which otherwise has regular initial stress on content words), were also tested and were found to fall in between the two ends along the stress deafness scale. The stress deafness of French speakers could also be confirmed using the same methodology within an L2 context, namely, French learners of Spanish (Dupoux et al., 2008).

Since the seminal work by Dupoux and colleagues on the perceptibility of stress contrasts, the extent of stress deafness and its typological repercussions have been investigated in detail with a wider range of stimuli on typologically distinct L1 populations. For example, Altmann (2006) expanded the focus on stress perception in L2 learning in an investigation of the ability to locate stress in English 2-, 3-, and 4-syllable nonce words by highly proficient learners from various typologically different L1s, which yielded two distinct clusters of L1s: (i) Speakers of L1s without word-level stress (Chinese, Korean, Japanese) and phonologically irregular stress (Spanish) displayed very successful perception abilities, as did the English native speakers, and (ii) speakers of L1s with phonologically regular stress (French, Turkish, and Arabic) performed significantly poorer. This, on the one hand, confirmed a relatively high degree of stress deafness for French speakers, on the other hand, Turkish and Arabic learners were not statistically different from the French group, although stress in these L1s is not as surface-observable as in French¹³, which poses some problems for the stress deafness hierarchy. Instead, these results were taken to provide support for the hierarchical *Stress Typology Model* (Altmann, 2006), where poor stress perception is attributed to an accumulation of positive settings for specific binary stress parameters. Since in this model non-stress languages do not incur any positive parameter settings for stress, speakers of these languages are expected to display good perception abilities (and, independently, tonal languages may also be more attuned to phonetic cues such as pitch or duration). Languages with regular phonological stress, however, incur a number of positive settings (specifically, for the parameters 'stress language', 'predictable', 'quantity sensitive', and 'left' or 'right' word edge), which are claimed to gradually impede the ability to perceive stress. French and Turkish would fall into the same category according to this model (both have predictable quantity insensitive stress placement at the right edge of the word) and are thus expected to behave similarly. Arabic, on the other hand, incurs one more positive setting since stress assignment in this language is quantity sensitive (all other parameter settings being identical with French and Turkish). Although it did not reach statistical significance, the Arabic learners' performance in this study was consistently lower than that of the other two predictable stress L1s.

Building on the findings in Altmann (2006), Kijak (2009) introduced additional criteria such as functional load of stress in the L1 and phonetic properties as potential factors that may influence the ability to perceive L2

stress. An investigation of L2 Polish, where stress is phonologically fixed on the penultimate syllable, with a variety of typologically different L1 groups and a design very similar to the one employed by Altmann, yielded four clusters of languages that did not differ significantly from one another within a given group: (i) Chinese and French; (ii) French, English, and Polish; (iii) English, Polish, Spanish/Italian, German, and Czech; (iv) Spanish/Italian, German, Czech, and Russian (groups ordered in ascending rate of success). Some of these findings were unexpected according to the Stress Typology Model and the stress deafness account. Specifically, these accounts cannot straightforwardly explain why the English and Chinese participants fared rather poorly (L1 stress not phonologically predictable and no word stress, respectively) while the Czech learners (L1 stress regularly on the initial syllable), especially in comparison with the other fixed L1 groups in this study (i.e. French and Polish), did surprisingly well. Kijak argues that the function of stress, rather than surface properties, in the L1 should serve as a predictor for L2 stress perceptibility, together with phonetic cues of stress in the L1. English native speakers have been found to make less use of stress in word recognition in their own L1 than German or Spanish speakers since vowel quality is already a good indicator for stressed syllables in English (e.g. Cutler and Pasveer, 2006). This may also affect their L2 stress perception ability, but requires further investigation since in previous L2 stress studies (in which English native speakers served as control group), they usually performed at ceiling level even for words that contained more than one full vowel (e.g. Pater, 1997; Altmann, 2006). A functional account may explain the outstanding performance of Czech learners compared to French ones. Kijak argues that Czech uses stress crucially for word segmentation, which is used to explain why Czech speakers are better at perceiving L2 stress than French speakers, for whom stress may be a more phrasal phenomenon and lack function at the word level. This, however, does not explain why Polish native speakers patterned with both Czech and French speakers. Furthermore, the claim that Czech speakers, but not French speakers, use word stress for word segmentation does not accord with recent psycholinguistic findings. In a word-spotting study, both French and Turkish speakers were found to use stress as a reliable cue for word segmentation (Kabak et al., to appear).

On the other hand, phonetic cues of stress in the target language are used to explain the comparatively poor performance by Chinese learners. In Polish, stress manifests itself mainly in a combination of high F₀ and a sharp F₀ slope rather than a more complex combination of several cues as in English, which supposedly makes it harder for tonal speakers to identify stress by relying on familiar phonetic properties that are also used for tone. It remains unclear, however, why Chinese speakers would fail to utilize pitch information in this case, which has indeed been shown to be the most reliable cue Chinese speakers use for stress perception (Wang, 2008).

3.2.3 Structural Factors in the Perception and Production of L2 Stress

What are target-language-internal factors in the acquisition of L2 stress? In a series of studies, Guion and collaborators (Guion et al., 2003; Guion et al., 2004; Guion, 2005; Wayland et al., 2006) investigated the effects of syllabic structure (e.g. light versus heavy syllables), lexical class (verbs versus nouns), and stress patterns of phonologically similar words on the production and perception of nonce bisyllabic English words by Spanish, Korean, and Thai learners, who were further divided into early (i.e. who started learning during childhood) and late learners (who started learning as adults), respectively. Overall, it can be stated for production that late learners, regardless of the L1, relied more on an extension of stress patterns of known words to novel words. However, they were less able to extract abstract generalizations across lexical classes or syllabic structures than the English control group, with early learners employing more similar strategies to the monolingual English control group than late learners. Transfer of L1 patterns could also be attested but only for late Spanish learners. Thus, there is a varying degree of success in the acquisition of different factors that arguably contribute to stress placement in English (e.g. word class, analogy, syllable structure), depending on, for example, differences in the respective prosodic systems or age of acquisition (Wayland et al., 2006).

In stress perception, English native controls were found to make use of information about lexical class, syllabic structure and phonological similarity to already existing words while L2 learners varied in how they utilized these factors. For example, early Spanish and early Korean learners employed all three, late Spanish learners relied mostly on syllabic structure, for late Korean learners the most influential factors were a combination of phonological similarity and syllabic structure, while for all Thai learners phonological similarity was the most contributing factor to successful stress perception.

Although it is difficult to make sensible generalizations and predictions from these findings for learners with different L1 backgrounds as well as for the acquisition of typologically distinct target languages other than English, it must be pointed out that such differences between groups indicate the importance of investigating not only the success rate but also the underlying strategies that lead to stress production and perception. Although learners might not use the same strategies as native speakers do, they may, however, still display native-like performance, as shown, for example, in the high rates of correctness in perception for Korean early and late learners in this study.

3.3 Summary

From such large-scale albeit restricted studies, it can be induced that poor perceptual abilities or stress deafness by no means imply that speakers who

display these characteristics are not able to perceive stress or stress contrasts at all. Instead, it is more conceivable to assume that these speakers have a significantly lower success rate than others, but are usually still well above chance level. As could be seen in the production studies (cf. Altmann, 2006; Kijak, 2009), learners of L1s with predictable stress (who usually show some degree of stress deafness) were able to yield, on the one hand, performance that did not correspond to their respective stress deafness rating and, on the other hand, more target-like productions than learners of non-stress-deaf L1s. This would not be possible without some (albeit restricted) ability to perceive stress in the L2. Thus, learners must be able to extract and create some abstract representation for at least a certain amount of information regarding L2 stress in order to utilize this for active stress assignment to nonce words in production. It is clear, however, since perfect perception does not automatically lead to perfect production (cf. Altmann above), that there is no straightforward connection between the perception and production of L2 stress. This is on a par with the observations of Guion and colleagues that target-like success rates alone do not necessarily imply target-like strategies. In addition, specific phonetic properties or the functional load of stress in the L1 may affect L2 stress perception and production as well (cf. Kijak, 2009).

The question of learnability in the course of L2 acquisition cannot be conclusively answered yet either. This hinges on the question if perception of stress is really an ability or rather an aptitude. We have seen tendencies of more advanced learners performing better than beginners, but these were never longitudinal studies that could reliably trace the development of individual speakers. At this point, it can be stated that L2 learners are affected by prosodic and phonetic properties of their L1 but that they may be able to employ other grammar-internal aspects and mechanisms of the target language to compensate, more likely regarding L2 performance rather than competence.

4. Conclusion

Our survey of various studies in the field of L2 phonology, a field still in its infancy, shows that developing L2 phonological grammars involves a highly complex enterprise with various structural and psycholinguistic factors at play. We highlighted the role of structural similarities and differences between the L1 and the L2, as well as universal tendencies, translated as 'markedness' into linguistic theorizing, to explain the peculiarities and variation observed in interlanguage phonological grammars across different L2 learner populations. Accurate perception of L2 contrasts, from the level of segments to suprasegmental elements, has been shown to play a primal role for successful acquisition of L2 contrasts. Experimental research has convincingly indicated how tenacious

L1 phonological knowledge can be in L2 perception. However, there is compelling and ever-growing evidence that accurate perception in L2 does not guarantee accurate production, which suggests that L2 speakers' phonological representations are not necessarily identical to those of native speakers. Therefore, it is not just whether the contrast is perceptible to the ear, but precisely what the contrast is and how it is implemented that seem to determine the native-like encoding of L2 phonological contrasts. It is then more sensible to characterize an L2 phonological system not as an imperfect imitation of native speech, but as the realization of imperfect representations. We have shown that L2 studies should look at both perception and production to make sensible generalizations about the nature of L2 representations. Needless to say, more cross-linguistic research, informed by both linguistic and psycholinguistic theory, is essential to fully understand the dynamics of L2 phonological acquisition. We hope to also have shown that research on L2 phonology is highly instructive in that it provides unique and invaluable opportunities to understand the nature of phonological systems, their internal organization and the dynamics of phonological processing.

5. Notes

1. We refer to common tendencies found in the phonological systems of languages as phonological universals (e.g. all languages have stops) although there seems to be only a handful of absolute phonological universals that are unequivocally grounded in synchrony rather than diachronic developments (see Hyman, 2008). Although controversial, several researchers have proposed parallels between phonological and syntactic universals, the latter of which is often assumed to be given by the Chomskian notion of Universal Grammar, an innate set of principles of language thought to be shared by all languages. The reader is referred to Carr (2006) on a recent discussion of the nature of universality in phonology and the relationship between phonological knowledge and innate linguistic knowledge.
2. Naturally, successful acquisition of a contrast also implies being able to adjust and fine-tune as well as retrieve a lexicalized contrast depending on contextual variation (e.g. in a sentential context; speech rate etc.) and environmental conditions (e.g. noise).
3. For age-related differences in SLA, see, for example, Bongaerts et al., (1997), Birdsong (2005), DeKeyser and Larson-Hall (2005), Muñoz (2008).
4. On phonetic perception in second language learners, the reader is referred to Strange (1995) and several other articles therein. Concerning intelligibility of L2 speech, there is a growing body of literature that suggests that the L1 backgrounds of talkers and listeners play an important role in determining the intelligibility of speech and that native listeners find native speech more intelligible than non-native speech (Munro, 1998; Munro and Derwing, 1999; Munro et al., 2006; Bradlow and Bent, 2002, 2008).
5. Universal Grammar (UG) is an innate set of generalizations about language that is assumed to underlie all human grammars. Assumptions on UG vary in different components of the grammar. For phonology it refers to a set of principles that govern

universal tendencies in phonological systems across the languages of the world (see below for concrete examples of phonological constraints that are attributable to UG in Optimality Theory; for the role of UG in the L2 acquisition of syntax, see Flynn, 1983; White, 1989; Schwartz and Sprouse, 1996).

6. Brown (1998) assumes that phonemes consist of distinctive features that are organized into a systematic hierarchy of constituents known as feature geometry (e.g. Clements, 1985). In this approach, a segmental representation contains only the information that is necessary to contrast it from all other segments in the system. Thus, the precise representation of a segment largely depends on which segments it contrasts with in a particular language (e.g. Avery and Rice, 1989).
7. Repetition priming refers to the assumption that an already studied stimulus (e.g. an object that has already been presented to the subject) influences the subsequent response to the same stimulus.
8. Following Selkirk (1982), Broselow and Finer (1991) assume the following sonority hierarchy, from the least sonorous to the most sonorous: Obstruents – Nasals – Liquids – Glides – Vowels. Each class of segments is then assigned a sonority value (Stops = 1, Fricatives = 2, Nasals = 3, Liquids = 4, and Glides = 5), which is used to calculate the sonority distance of permissible clusters.
9. The status of stop+glide sequences is controversial. For example, Sohn (1987) argues that Korean glides are part of the nucleus while Lee (1982) supports the idea that glides form the onset together with the preceding stop.
10. For earlier accounts of interlanguage phonologies, see Eckman (1981) and various articles published in Ioup and Weinberg's (1987) seminal volume on interlanguage phonology.
11. Optimality Theory (Prince and Smolensky, 2004 [1993]; McCarthy and Prince, 1995) proposes that observed surface patterns in languages are the result of the interaction among conflicting constraints. Constraints are assumed to be universal and are of two types: (i) faithfulness constraints, which ensure that input representations and output forms are identical, and (ii) markedness constraints, which assess how well-formed output forms are in light of markedness considerations. Different constraint rankings give rise to different grammars. Accordingly, second language acquisition is assumed to be the process of readjusting an already existing constraint ranking (i.e. the L1 grammar).
12. In an AX discrimination task, the subject is asked to determine whether two stimuli separated by an interstimulus interval (e.g. 1500 milliseconds) within a trial are the same or different. In such tasks, stimulus A may vary (categorically or gradually) from X, with expected responses of 'same' or 'different'.
13. In French, word stress will always be at the right edge of a word (the only exception being final schwa syllables), while Turkish stresses the right edge of every phonological word (which may be followed by unstressed suffixes and thus not immediately coincide with an utterance edge; cf. Kabak and Vogel, 2001 for details). Furthermore, there are exceptions to the regular stress pattern in Turkish, for example, place names. Arabic has regular quantity-sensitive stress placement. In simple terms, it falls on the rightmost non-final heavy syllable (see McCarthy, 1979a or Hayes, 1995b for details). Word stress would thus only directly coincide with utterance boundaries in the case of superheavy final syllables, which would receive final stress, otherwise it would always be placed at least one syllable from an utterance edge.

13

Phonological Disorders

Dirk-Bart den Ouden

Chapter Overview

1. Introduction	320
2. Phonological Impairments in Aphasia	321
3. A Brief History	325
4. A Closer Look at Phonological Structure in Aphasia	332
5. A Two-Case Study of Syllable Structure Effects on Lexical and Postlexical Errors	334
6. Conclusion: Formal Phonology and Aphasia	340

1. Introduction

This chapter provides an overview of research in the field of phonological impairments in aphasia, with an emphasis on phonological output deficits and on issues that pertain specifically to phonological *structure*, rather than the mere presence or absence of word-form errors, or the relation between phonological and semantic or syntactic representations in the lexicon. I will take a global look at the different directions research on phonological structure in aphasia has taken in the last 50 years and discuss remaining issues along the way. Particular attention is paid to the different perspectives researchers have taken in their investigation of phonological impairments, namely by using phonological and psycholinguistic theories to inform the analysis of impaired phonological processing, or the other way around, by letting phonological impairments inform ideas about psycholinguistic theory and phonological theory, structure and representation. The chapter ends with a two-case study investigating syllable structure constraints on word-form representations and accessibility at two different levels of phonological processing.

For a more general overview of clinical applications of phonological theory, the reader is referred to the recently published textbook by Ball et al. (2010), which does reflect the heavy emphasis in the clinical phonology literature on developmental language disorders, over the neurogenic language disorders that are the topic of the current chapter.

2. Phonological Impairments in Aphasia

Aphasia is an acquired impairment of language in adults, resulting from focal brain damage. Such damage may affect different language modules and modalities, so different forms of aphasia are not regarded as communally stemming from the same deficient functional level of processing. Even within the group of speakers with phonological output defects – the focus of this chapter – different speech patterns can be recognized, sometimes implying differently affected parts or stages of the phonological system. Brain lesions are not constrained by functional module boundaries and individual patients may differ greatly in their patterns of aphasic symptoms. Nevertheless, it is possible to distinguish clusters of symptoms in aphasia, regarded here as ‘syndromes’, which are hypothesized to result from different functional loci of impairment, even if the particular clustering of symptoms into a syndrome may well be more the result of anatomical factors than of functional associations (cf. Poeck, 1983).

Many aphasic speakers present with so-called *literal paraphasias* in production. This term covers word-form errors from which the target word is still recognizable, but in which sounds are substituted, adapted, omitted, transposed or added, and it excludes neologisms and verbal or semantic paraphasias. It is preferred, here, over the terms *segmental* or *phonemic paraphasia*, which may also be encountered in the literature, because it is by no means always clear whether the erroneous units in such paraphasias are ‘phonemes’, rather than the phonological or phonetic features these phonemes consist of (cf. Harris, 1998), or even whether they are caused by faulty articulatory planning or execution directly (Ball et al., 2010). For example, in English, a production of palatal fricative [ç] for target alveolar /s/ would generally be considered a phonetic distortion, as /ç/ is not a part of the English phoneme inventory. It would be odd, then, to consider production of post-alveolar fricative [ʃ] for target /s/ as a phonemic error, simply because /ʃ/ is part of the regular English system (Ball et al., 2010), while the two errors clearly have a lot in common. Rather, the nature of these errors, that is, segment substitutions or phonetic distortions (e.g. ‘backing’) should be subject to investigation, pending which it is better to use a term that is not biased towards one interpretation.

The three main aphasic syndromes in which large proportions of literal paraphasias generally occur in speech output are Wernicke’s aphasia, conduction

aphasia and Broca's aphasia. Wernicke's aphasia is a lexical-semantic deficit, with impaired comprehension. Production is fluent, with natural prosody, and characterized by semantic and/or literal paraphasias. In more severe cases, speakers with Wernicke's aphasia may produce many neologisms. They may also present with *logorrhea*, in which case their continuous speech stream is difficult to interrupt, combined with an apparent lack of error-awareness.

For conduction aphasia, production of literal paraphasias and repeated attempts at error correction are the central characteristics. These repeated attempts at error correction are known as *conduites d'approche* (Goodglass and Kaplan, 1983), or *phonemic approximations*. Patients are generally fluent, but speech may become hesitant and dysfluent in the face of phonological errors, of which the patient is well aware. Comprehension is relatively intact. A distinction must be made between 'repetition conduction aphasia' and 'reproduction conduction aphasia'. The first is the form of conduction aphasia classically conceptualized as the result of a disconnection between auditory and motor speech representations, manifesting itself solely in impaired auditory repetition, especially of nonwords (Wernicke, 1874[1977]; Lichtheim, 1885; Geschwind, 1965). Although the term itself is not particularly transparent, 'reproduction conduction aphasia' indicates the deficit in which the phonological output problems are not limited to repetition tasks, but also occur in spontaneous speech and in naming tasks, to much the same degree (Caplan and Waters, 1992). Reproduction conduction aphasia is in fact the disorder most associated with deficient application of phonological output processes (cf. Den Ouden and Bastiaanse, 2005). In what follows, this is the deficit that is implied when the term 'conduction aphasia' is used without further specification.

Broca's aphasia is an output deficit marked by nonfluent speech and problems with grammatical morphemes and sentence building. Impaired comprehension of complex syntactic structures has also been shown for this category of aphasic patients (e.g. Caramazza and Zurif, 1976; Berndt et al., 1996). The literal paraphasias produced by Broca's aphasics have been claimed to result from 'phonetic disintegration' (Alajouanine et al., 1939), that is, the inability to translate a correct phonological speech plan into its correct phonetic counterpart and articulation. These speech errors are considered to be generated at a more peripheral, phonetic level than the paraphasias of the other two syndromes (Blumstein, 1991). Broca's aphasia often coincides with diagnosed apraxia of speech (AOS) and, indeed, many researchers treat the phonological output problems in Broca's aphasia as equivalent to those in isolated AOS. From that perspective, that is, strictly within the phonological domain, Broca's aphasia is the same as AOS. It is likely, however, that the phonological problems that often occur in typical Broca's aphasia are caused by a combination of AOS and the kind of phonological disruption that occurs in a more pure form in conduction aphasia. Severe AOS may mask any additional phonological problems in speech

production, but a purely articulatory (planning) deficit should be limited to the speech module and not be present in written output, whereas this is not the case for phonological impairments at deeper levels of processing (though see Visch-Brink, 1999, for a dissociation between graphemic and phonological representations).

In a simple modular representation of the speech production process, distinguishing lexical representations and retrieval from phonological encoding and articulatory planning modules (as in Figure 13.1), the functional deficit of conduction aphasics would be located at the level of phonological encoding (e.g. Kohn, 1988). These speakers mainly have (postlexical) problems with the sequencing of sounds and sound patterns from the lexicon. Wernicke's aphasics' phonological problems are caused by deficient retrieval from the lexicon, or even by distorted lexical representations themselves. In Broca's aphasia, the main deficit would be in the articulatory planning stage, in such a model.

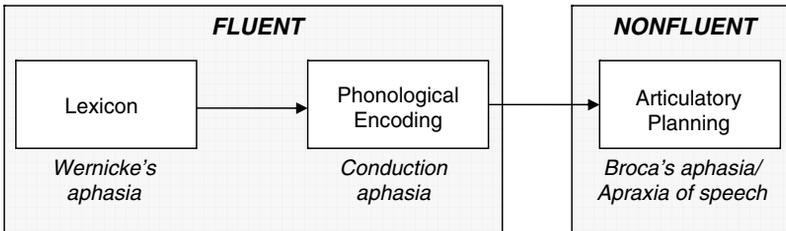


Figure 13.1 An idealized model of functional loci of the origins of literal paraphasias

It is important to note that, although this schematic modular representation of the way that different phonological output disorders are conceived is widely used, formally or informally, most researchers will also grant that the model is inadequate for the representation of a particular functional deficit location in any specific individual patient. Most, if not all, aphasic speakers will reveal the system of language representation or the extent to which it is impaired to be more gradient and less focal than this modular representation suggests. It is precisely this perceived gradience that has been a much-debated topic within research into phonological deficits in aphasia. As an idealized research model, however, it serves as a useful starting point for many investigators.

One way of dealing with the near-absence of 'clean' aphasic functional deficits is to generalize over larger groups of aphasic speakers that share a particular characteristic of impaired speech output, under the assumption that more random individual differences between subjects will be cancelled out in group analyses. The most widely used division, in this respect, is that between fluent and nonfluent aphasic speakers, as also shown in Figure 13.1.

Nonfluent speakers, then, are considered to have a deficit in articulatory planning, while fluent speakers have a problem in the derivation of correct phonological output forms, either caused by deficient (access to) representations, or due to deficient online application of phonological processes in production. Importantly, the nonfluent group excludes speakers with dysarthria, who have a direct articulatory impairment caused by loss of motor control over specific articulators, or even paralysis of articulators. Such deficits are not deemed to be of a linguistic nature, and therefore not covered by the term 'aphasia' as commonly employed, while the extent to which apraxia of speech is to be considered a 'linguistically' defined deficit is more a matter of research and debate (cf. Code, 1998; Dogil and Mayer, 1998). The division between fluent and nonfluent aphasia, thus defined, has allowed researchers to investigate phonological versus (cognitive) phonetic impairments. An analogous distinction is that between 'anterior' and 'posterior' patients (cf. Blumstein, 1991). The latter categorization refers to lesion sites in aphasic patients, either in the left frontal lobe (anterior, nonfluent), or in the left temporal lobe or temporoparietal area (posterior, fluent).

The view that phonological output problems may stem from deficits at different levels of processing, and may thus reflect different aspects of phonological processing, has evolved over time, concurrent with the evolving of phonological theory itself. Rather than providing a complete overview of all perspectives on aphasia and phonological deficits in what follows, I will focus on approaches that make reference to phonological structure, representation, rules, derivations or constraints on output well-formedness. For this reason, approaches that describe phonological deficits in terms of short-term-memory disorders, as suggested first by Warrington and Shallice (1969) (see also Wright and Shisler, 2005), are not discussed here. However, this absence should not be taken to imply that such analyses are not relevant to or incompatible with accounts that make direct reference to phonological structure, only that they do not speak directly to structural systematicity in the phonology of aphasia. The view adopted here is that even if the underlying deficit in a particular type of phonological disorder is one of limited short-term memory capacity, or impaired articulatory rehearsal, as has been argued, for example, for repetition conduction aphasia (Shallice and Warrington, 1977) and apraxia of speech (Rochon et al., 1990), phonological systematicity in output forms and errors, if present, must still be described and accounted for in terms of phonological theory.

One important issue to address at this point is the difference between the nature and the objectives of formal phonological theories, psycholinguistic theories and accounts of language impairment. The main goal of formal phonological theories is to capture linguistic generalizations in either a computationally or representationally economic manner. Psycholinguistic theories are generally concerned with models of the (time) steps and the processes through which

language is produced and comprehended and with models of how our minds organize linguistic representations, such as lexical items. Neither of these make direct claims about the other and in fact many scholars will argue that the performance data that often form the basis of psycholinguistic theorizing and testing cannot be used to falsify formal linguistic theories, which typically aim to characterize language competence (cf. Chomsky, 1972; McCarthy, 2002). Similarly, formal linguistic arguments are claimed to bear no relevance to psycholinguistics. Data from neurogenic language deficits, then, are found to be even less relevant to either formal linguistic or psycholinguistic theories. In the aphasiological literature, however, both formal linguistic and psycholinguistic theories are commonly used to inform accounts of performance patterns in language deficits. The approach taken in the present chapter is that this is rightly so. Ultimately, any theory about human behaviour or characteristics can only approach relevance if it can incorporate or be compatible with how humans actually 'work', in this case how their brains and physical speech apparatus support language representation and use. Not every theory needs to address all these issues, that is, a formal phonological theory may be restricted to part of the system, but in the end, theories describing different modules, or modalities, have to fit onto one another in order to come as close as possible to reflecting 'reality'. As such, a phonological theory does not have to be about performance, but it cannot be *incompatible* with performance. Note that the test of compatibility may well make reference to a separate theory mapping language performance onto perhaps more abstract linguistic representations and processes. This also implies that insights may be gained from discussing patterns of phonological output deficits in terms of formal phonological theories, in terms of psycholinguistic production models and in combinations of these, as long as the scientist is aware of the different components she is bringing together. The reader, in turn, should be aware that this represents a particular standpoint taken as part of an ongoing debate within the broader field.

3. A Brief History

Roman Jakobson (1941) famously hypothesized that aphasia was the mirror of child language acquisition, in that language elements that were acquired last, that is, 'marked' elements, would be first affected in language loss due to stroke. The generative phonology approach introduced in *The Sound Pattern of English* (SPE) (Chomsky and Halle, 1968) placed an emphasis on phonemes as independent feature bundles, without further structure within the system of phonological features. Syllable structure was also absent from SPE, in which a single level of analysis, with input-to-output-mapping rules applying serially, was used to describe and derive (i.e. generate) phonological regularities.

The reigning perspective in the late 1960s, then, was that across different types of aphasia, patterns of phonological output errors would largely be uniform. Within the phonological system, no separate levels of processing that might be independently affected were conceived in the contemporary psycholinguistic models, and the prediction was that inherently marked elements would be 'first out' in aphasia.

Lecours and Lhermitte (1969) investigated systematicity, in terms of phonetic feature changes, in the literal paraphasias of speakers with 'phonemic jargon aphasia', roughly equivalent to Wernicke's aphasia, that is, speakers whose output contained no elements of what Lecours and Lhermitte (1969: 194) call 'phonetic disintegration'. Phonemes were modelled on a 3-dimensional grid, which directly represented the distance between different phonemes in terms of their shared phonetic features. Lecours and Lhermitte's (1969: 201) finding was that the paraphasias of the jargon aphasic speakers showed a 'definite tendency to replacement of phonemic units by morphologically similar ones'. In other words, the literal paraphasias of speakers with a phonological deficit were shown to be driven by phonetic distance, an apparently physiological measure.

Blumstein's (1973) starting hypothesis was very Jakobsonian, in that she appears to have expected to find, and indeed did find, similar effects of the relative phonological markedness of segments in patients with different types of aphasic syndromes: '[Regardless] of the area of brain damage, the more complex phonological structures are impaired and the less complex phonological structures are relatively preserved' (Blumstein, 1973: 136). Blumstein studied the error patterns in the speech of patients with Broca's aphasia, conduction aphasia and Wernicke's aphasia. Patients with these three types of aphasia all made more errors on segments that had been classified as more marked beforehand. In segment substitutions, marked segments were generally replaced by less marked segments.

Using motor complexity as opposed to phonological markedness as the gradient scale on which to classify segments and segment types, Trost and Canter (1974) and Burns and Canter (1977) also found comparable error patterns in similar groups of patients to those studied by Blumstein. All groups made more errors with increasing motoric complexity, from vowels to singleton consonants, to consonant clusters, the latter yielding most errors in all groups. However, Burns and Canter (1977: 501) note that motoric simplification was not the sole driving force behind paraphasic errors: 'While errors involving more complex productions than required for target were common to both [Wernicke's and conduction aphasic] groups, the pattern was particularly pronounced among the Wernicke's aphasics'.

One issue in the analysis of speech errors is that human categorical perception ensures that subtle phonetic distortions in sound production will still be categorized as phonemes known to our language system. Therefore, a mere

perceptual analysis of literal paraphasias will not be able to differentiate adequately between paraphasias of a phonemic nature and those of a subtle phonetic nature (see Mowrey and MacKay, 1990; Wood and Hardcastle, 2000). Blumstein et al. (1977, 1980) performed acoustic analyses to investigate the production of voice onset time (VOT) in Broca's, Wernicke's and conduction aphasic speakers of English. The (partially) voiced English plosives have a shorter VOT range than voiceless plosives and these ranges typically do not overlap in normal production. However, Broca's aphasic speakers turned out to produce many speech sounds with deviant VOTs, outside the normal range. Such errors were interpreted to stem from a disorder in articulatory timing. Importantly, even the Wernicke's aphasics showed some instances of deviant VOT production, and the number of deviant VOTs produced by conduction aphasics was found to be in between that of the Broca's and the Wernicke's. The boundary between the articulatory-timing disorder and more phonemic disorders could not be drawn sharply, therefore, suggesting gradience in the types of phonological disorders, rather than strict modularity.

Gradience may also be deduced from the fact that error types are not exclusive to speakers with particular phonological output disorders. Conduction aphasic speakers, for example, are not alone in their production of *conduites d'approche*, although it is a prime characteristic. Joannette et al. (1980) analysed such phonemic approximations in Broca's conduction and Wernicke's aphasic speakers and found that speakers in the first two groups appear to retain stronger knowledge of the target word forms. While the sequences produced by Wernicke's aphasics typically strayed further and further from the target, those produced by Broca's and conduction aphasic speakers remained closer to the target word-form, or gradually approached it, even if this did not always lead to successful production in the end.

Following work such as by Fudge (1969) and Kahn (1976), syllables came to be recognized as domains for phonological processes in the 1970s, adding a representational layer to the SPE perspective (see also Szigetvári, this volume). By the early 1980s, internal syllable-structural representations were regularly incorporated in phonological analyses (e.g. Selkirk, 1982), including analyses of impaired speech. Blumstein (1978: 194) noted that consonants within tautosyllabic clusters are less prone to substitution errors than single consonants in aphasia, arguably because they are highly interdependent and have less 'structural independence', as also reflected in coarticulation. Studies by Nespoulous et al. (1984, 1987) took into account syllable structure in the analysis of literal paraphasias produced on single-word reading and repetition tasks by Broca's and conduction aphasic speakers. They found that the error patterns of Broca's aphasics were generally constrained by the relative markedness of different segments and syllable structures, whereas such constraints were not found in the errors of conduction aphasics. Broca's aphasic substitution errors were

closer to the target phonemes in terms of feature changes, and they would, for example, show a higher proportion of devoicing of voiced plosives, voicing in plosives being a marked feature. Consonant clusters would often be reduced to single segments in Broca's aphasics' paraphasias. Conduction aphasic speakers showed greater variability in terms of the 'markedness direction' of their paraphasias, which notably contained a higher proportion of transpositions and other 'sequential' errors, and contained almost as many cluster creations as cluster reductions. These speakers also produced relatively more errors on longer words. They were concluded to suffer mainly from a serial-ordering deficit, at the level of phonological planning.

Importantly, Nespoulous et al. (1984: 213) point out that '[if], as can be reasonably claimed, motor simplicity and complexity closely parallel abstract phonological simplicity and complexity, the task of distinguishing the respective roles of both potential determinisms in Broca's aphasics' error production might turn out to be a bottomless chasm and, conceivably, an irrelevant venture'. This speaks directly to the debate between functionalist and more nativist approaches to formal phonological theory; if all regularities can be traced back to motor (or perceptual) complexity, there seems to be no reason for positing additional underlying structure to account for these same regularities, in acquisition patterns, language change, typology or for reasons of learnability (for discussion, see contributions to Burton-Roberts et al., 2000). For Nespoulous et al. (1984, 1987), therefore, the psycholinguistic model, in which the conduction aphasic deficit is at an abstract level of phonological representation and Broca's aphasic speakers have a phonetic planning deficit, informs conclusions about the nature of phonological representation at the abstract level. Their conclusion is that the constraints that yield markedness effects are not present at this phonological planning level and that it is 'the motor aspects of markedness which produce the effect' in Broca's aphasia (Nespoulous et al., 1984: 219).

In a summary of her own work and that of others, Kohn (1988) presented the arguments for employing the psycholinguistic model as presented in Figure 13.1, with three separate deficits mapped onto different stages in the phonological production process. Wernicke's aphasic speakers were deemed to have a lexical (access) deficit, reflected by errors that add structure (in terms of CV structure as well as number of syllables) and the fact that they produce the smallest number of target phonemes. Conduction aphasic speakers, then, had a sequencing deficit, reflected by production of incomplete strings of target phonemes, the use of ordering strategies, such as overt spelling during naming tasks, and by their production of *conduites d'approche*. The deficit in Broca's aphasia was at the level of articulatory planning, reflected by motoric simplification in cluster reduction, that is, production of many unmarked syllable structures (CV), far fewer errors on vowels than on consonants, and few errors that changed syllable structure position of segments, as in transpositions. Kohn (1988) goes

on to propose 'a neurolinguistic recasting of the phonological production model' (p. 114), in which a left-hemisphere serially operating network supports phonological lexical access (posterior superior temporal area), construction of phonemic strings (inferior parietal area) and integration of context-sensitive phonetic information (posterior inferior frontal area), in a serial feedforward fashion, much as in the network proposed later by Hickok (2001).

This modular representation of aphasia types has been highly influential since the end of the 1980s, also fed by the publication of Levelt's psycholinguistic model of unimpaired speech production (Levelt, 1989, and later Levelt et al., 1999), onto which these types could be mapped transparently. However, there also remained the sense that real data, particularly from individual case studies, showed that the difference between types of aphasia, even restricted to the phonological domain, were often of a more gradient nature than allowed by this modularity. Two related questions remained: (i) Are aphasic syndromes restricted to their own modules? (ii) Are phonological 'rules'/processes restricted to specific modules? Particularly the second question shows how the modular model (as in Figure 13.1) invites the use of aphasic syndrome characteristics as windows on its individual components, that is, levels of processing.

The likely interaction between phonological and phonetic levels of representation is pointed out by Dogil and Mayer (1998), who claim that AOS is really a phonologically defined deficit, rather than solely a problem with articulatory planning. Dogil and Mayer (1998: 55) present data of speakers with AOS who produce relatively more errors on 'phonologically underspecified and phonetically transparent' speech units, such as schwas, and fewer errors on highly specified segments, such as Xhosa clicks. Stated in terms of phonological lexical underspecification theory (Archangeli, 1988), it is argued that the problem in AOS is one of overspecification of speech sounds at an underlying level, which is responsible for the noted lack of coarticulation in apraxic speech, as well as for greater problems with elements that should be *least* marked.

Christman (1992) investigated the production of neologisms by speakers with different types of aphasia. Her analysis is in terms of sonority distribution within syllables, where sonority should rise from syllable margins to the syllable peak, according to the Sonority Sequencing Principle (Jespersen, 1904). The fact that neologisms in different types of aphasia do not violate this principle, that is, no output strings like **rpitn* were produced, leads Christman (1992: 70) to conclude that '[sonority] may indeed be a well-distributed aspect of language knowledge to which computational mechanisms at different levels have access during language production'. Closely related to this are earlier observations that phonological errors in aphasia do not seem to violate language-specific phonotactics – they overwhelmingly adhere to the well-formedness rules of the aphasic speaker's native language (Buckingham, 1987). The often noted hierarchy in substitution errors, where place features seem to be most susceptible

to change, followed by voice and manner (e.g. Southwood and Chatterjee, 1998), may also follow directly from a tendency to preserve syllable structure and language-specific phonotactics – place features of tautosyllabic cluster consonants can be changed far more easily than voice or manner features, without necessarily breaking up target syllable structure.

Wilshire and Nespoulous (1997, 2003) investigated patients with a deficit at the phonological planning (encoding) level, to get at the question whether syllables are represented as ‘chunks’ in the lexicon, including their segments, as proposed by Dell (1986, 1988), or whether they are built during a process of syllabification by which segments are mapped onto empty syllable structure nodes, as proposed by Levelt et al. (1999). Wilshire and Nespoulous’ (2003: 428) argument is that ‘[if] such patients’ performance can be shown to be sensitive to syllabic variables, such as syllabic structure or syllable frequency, then this would support the notion that the syllable is an important unit of processing during phonological encoding’. Wilshire and Nespoulous did not find an effect of syllable frequency and concluded that syllables do not play a role as ‘chunks’ at the level of phonological encoding, though possibly as frames, for the organization of segmental content.

Interestingly, Levelt’s proposal of a mental syllabary (Levelt and Wheeldon, 1994; Levelt et al., 1999; Cholin et al., 2006), a store from which syllable-sized gestures can be retrieved at the level of phonetic planning (as opposed to phonological planning), has also been used as the basis for accounts of AOS as a phonetic planning disorder. Varley and Whiteside (2001) propose that speakers with AOS have lost access to this store, which is hypothetically used only for high-frequent gestures (although Varley and Whiteside speak of word-size representations, rather than syllables). As a result of the syllabary deficit, speakers with AOS have to fall back on the segment-by-segment syllable assimilation system, which is usually only at play in novel or low-frequency syllable production. Apraxic speech errors, then, result from usage overload of this route. The main evidence for the proposal of a syllabary disorder in AOS is the absence of a syllable frequency effect on the error rates of speakers with AOS. Other studies, however, do show such frequency effects in AOS (Aichert and Ziegler, 2004), so the question of whether AOS may be considered a deficit in a mental syllabary has not been resolved and is still subject to debate. For an overview of arguments and experimental evidence related to the syllabary, see Cholin (2008).

While Varley and Whiteside use a psycholinguistic model and a hypothesis about the level of deficit in AOS to gain insight into the nature of that deficit, Wilshire and Nespoulous’ (1997, 2003) approach also starts from a deficit-level hypothesis, but is aimed at gaining understanding of phonological representation at this particular level. The construction of this argument is opposite to that employed earlier by Béland et al. (1990), who used a similar three-level model

to locate the functional level of deficit in a case study, consisting of Underlying Representations (UR), Lexical Representations (LR) and Surface Representations (SR). Béland et al. start from the position that syllable structure is not yet present at UR, but is assigned only at LR. Since their subject's substitution errors never force resyllabification, the conclusion is that the functional locus of deficit in this case must be at LR, the level of lexical representations.

Gilbers et al. (1997) specifically test a formal theoretical hypothesis by referring to phonological versus phonetic psycholinguistic levels of deficit. Based on acoustic evidence, they argue that phonological length has no phonetic equivalent, but that it is only an abstract representational (phonological) feature. Therefore, patients with a phonetic planning deficit (Broca's aphasic speakers), would not be expected to make phonological length errors, for example, producing [ɪ] for /i/, while this would be possible in speakers with a phonological output deficit (conduction and Wernicke's aphasic speakers). This is indeed what Gilbers et al. find. In this way, then, a psycholinguistic production model, coupled with hypotheses about levels of deficit, is used to inform formal phonological theory – in this case an autosegmental approach to phonological representations – where 'long' vowels occupy two positions on an abstract representational tier, which does not map directly onto phonetic duration (Goldsmith, 1979a; Clements, 1985).

As becomes clear from these lines of argumentation, aphasiologists must be wary of making their reasoning circular. It will not lead to further insight if we first build psycholinguistic models based on aphasic datasets, then take these models as *a priori* true, and subsequently map the same aphasic datasets onto these models, to make the models inform our conception of aphasia. Starting assumptions and hypotheses, therefore, have to be laid out transparently and motivated independently. Nevertheless, to the extent that aphasic symptoms can be associated with deficits at specific psycholinguistic levels, they can provide an even better window on underlying structure and processes than normal speech errors, which occur at relatively low frequency and may originate at any level of processing, without any external information about the likelihood of the specific origin of an individual error.

Of course, if all error characteristics are ultimately shown to be alike, this argues in favour of more holistic viewpoints, for example, the view that all literal paraphasias are generated at the same functional level, or the view that phonological constraints are pervasive throughout the speech processing system, hard-wired from the abstract to the concrete. In particular, Dell et al.'s (1997) proposal that different aphasic symptoms stem from the different settings of activation parameters that apply to the entire production system seems to be more suitable to account for 'holistic' data than for evidence of qualitatively different types of error patterns. However, their model of (impaired) language production may well be complemented with hierarchical phonological

structure. Such structure might reflect positional markedness effects within syllables, and in this way be able to account for the characteristics of different types of aphasia, if indeed these are found to be qualitatively different.

4. A Closer Look at Phonological Structure in Aphasia

A number of aphasiological studies have looked in greater detail at the manipulation of syllable structure, that is, the structure within syllables, in phonological output errors (Béland et al., 1990; Favreau et al., 1990; Valdois, 1990; Béland and Paradis, 1997; Romani and Calabrese, 1998). Valdois takes into account hierarchical syllable-internal structure in the analysis of aphasic speech errors on consonant clusters, arguing that the most adequate and insightful description is one along the lines of Government Phonology (Kaye et al., 1990), in which segments (phonemes) in governed positions are hypothesized to be more prone to error than those in governing positions. Valdois shows that the theoretically weaker positions are indeed vulnerable to deletion and substitution. This contrasts with a linear hypothesis according to which, say, the second position in each consonant cluster would always be most vulnerable, regardless of segment type. Valdois points out the likely interaction between the segmental and the syllabic levels of phonological representation: although omission errors were not strongly determined by phoneme type, neither were they completely phoneme-type-independent.

If consonant clusters are phonologically 'marked', or simply difficult to produce, speakers can break up such clusters either by deleting one element, or by inserting an epenthetic vowel. Béland and Paradis (1997) present an account in precisely those terms, named 'constraints and repairs'. A Preservation Principle aims to conserve existing segments, resulting in a preference for segment insertion when confronted with complex clusters. However, such insertion does entail adaptation of the existing structure, so it is costly as well. If the repair becomes too costly, the Preservation Principle may be overridden and structure may simply be reduced by segment deletion. In a case study of a speaker with primary progressive aphasia, followed over time, Béland and Paradis show an increase in the number of deletions, reflecting the patient's lowering tolerance threshold to more complex repairs. With respect to the type of material that is 'inserted' in epenthetic consonant cluster resolution, Buchwald et al. (2007) have argued that the vowels that were inserted by the aphasic speaker in their study (who had a phonological output disorder) were discrete phonological units, rather than the mere product of articulatory mistiming, as the insertions were qualitatively identical to lexical vowels.

Romani and Calabrese (1998) provide a detailed analysis of deletions, substitutions and additions in the literal paraphasias of an Italian aphasic speaker

with a combination of a deficit in articulatory planning and reduced verbal short-term memory, as assessed with a digit span task. Again, errors in syllable onset clusters were highly constrained by syllable-internal structure, analysed in terms of the Sonority Dispersion Principle (SDP) (Clements 1990), which states that syllables prefer a steep rise in sonority from onset to peak, and minimal difference in sonority between peak and coda. Interestingly, their subject showed very few reductions of s-obstruent clusters, which is in line with the analysis of these clusters as heterosyllabic (e.g. /s/ being extrametrical) and cluster reduction being triggered by complex syllable constituents, rather than mere consonant adjacency. Due to the phonotactic restrictions of the Italian language, Romani and Calabrese's study was confined to the analysis of errors in syllable onset clusters, in the absence of consonant clusters in syllable codas. Romani and Galluzzi (2005) compared repetition errors in a group study of Italian patients with phonological and articulatory deficits, showing that effects of syllable structure complexity were confined to the group of speakers with articulatory difficulties.

Den Ouden (2004) investigated effects of positional markedness on onset and coda cluster reduction errors in groups of fluent versus nonfluent aphasic speakers of Dutch, arguing that the dichotomy here was between patients with an inherently phonological (output) disorder and those with a phonetic planning disorder. His analysis, in terms of Optimality Theory (Prince and Smolensky, 2004 [1993]), allows not only for an interaction, but for a direct conflict between syllable structure preferences and preferences for segment type. As formalized in the SDP, syllables prefer to have a maximally non-sonorant onset and a sonorant rhyme, so that coda sonorants are preferred to coda obstruents (i.e. *pin* rather than *pit*). However, going back to Jakobson (1941), segmental markedness, independent of prosodic structure, is argued to prefer consonants to be maximally non-sonorant, and vowels maximally sonorant. These two markedness hierarchies, therefore, yield the same segment preferences in syllable onsets (non-sonorant consonants), but conflicting preferences in syllable codas (sonorant versus non-sonorant consonants). The nonfluent aphasic speakers in this study deleted sonorant consonants in both onset and coda clusters (*print* → *pit*), while the fluent speakers also deleted sonorant consonants in onset clusters, but showed equal numbers of sonorant versus non-sonorant deletions in coda cluster reduction (*print* → *pit*; *pin*). According to Den Ouden (2004), this suggests a dominant role for segmental markedness constraints in nonfluent aphasic speakers' grammars, and an overlapping ranking of segmental and syllable markedness constraints in the grammars of fluent aphasic speakers, resulting in the 50–50 error distribution in codas. Rather than positing a large-scale reranking of constraints between the two grammars, which would overgenerate possible error patterns, Den Ouden argues that some constraints may be at work only at specific psycholinguistic levels of processing (e.g. syllable

markedness), whereas others are present at all levels (e.g. segmental markedness). A deficit at a particular functional level is then represented as a lowering of faithfulness constraints at that level, sometimes resulting in the 'emergence of the unranked', as in the fluent aphasic speakers' error patterns.

Most studies have limited their investigations of literal paraphasias to the single-word level and the aphasiological literature shows that literal paraphasias can reveal the influence of phonological structure within syllables as well as within phonemes, or feature bundles. However, evidence of the influence of higher suprasegmental prosodic structure on literal paraphasias has also been shown, for example, by Bastiaanse et al. (1994), who present a case study of a conduction aphasic speaker who lengthens sonorous phonemes primarily at initial and final intonational phrase position, as if to 'highlight phonological structure' (Bastiaanse et al., 1994: 40). If phonological phrases serve as planning domains at the level of phonological encoding (Levelt, 1989), it might indeed be expected that patients with a deficit at this processing level show effects of phonological phrase length on their output errors (cf. Kohn, 1989; Den Ouden and Bastiaanse, 2005). Nevertheless, Pate et al. (1987) show that this was not the case in a conduction aphasic speaker's errors, which were not influenced by the size of phonological phrases, but rather by word length. They suggest that their patient plans segmental content over a restricted span, relative to that applied in unimpaired speech, possibly as a strategy to deal with the phonological planning or sequencing impairment itself. The disconnection between phoneme selection and suprasyllabic prosody is emphasized by Hanlon and Edmondson (1996), who present a speaker with 'phonemic jargon aphasia' and intact phonological phrasing, as exemplified by preserved intonation contours.

An interesting dissociation has also been observed between the likelihood of errors on vowels versus consonants, indicating either separate representations for these two classes, or distinct processing mechanisms, possibly related to syllabic constituency. In the majority of studies, consonants are reported to be more vulnerable to error than vowels (Canter et al., 1985; Béland et al., 1990), though case studies exist in which the reverse pattern is shown (e.g. Semenza et al., 2007; Caramazza et al., 2000). The first finding is more easily accounted for in terms of the prosodic prominence held by vowels in their peak syllable positions, but the latter observations do suggest a true double dissociation.

5. A Two-Case Study of Syllable Structure Effects on Lexical and Postlexical Errors

This section presents two cases of Dutch fluent aphasic speakers, who were both tested with a battery in which the same lexical items were tested in three different tasks, naming, repetition and productive phoneme monitoring.

The same items are used for optimal comparison between error rates on the different tasks (Butterworth, 1992). Items are further categorized with respect to length and phonological structure (complexity), to allow systematic investigation of the influence of these factors on error rates. The phoneme detection task, explained below, allows error analysis for effects of positional markedness. This is used to shed light on the amount of structure present at the respective patients' functional levels of deficit. The data and tasks presented here were part of a multiple case study, first reported in Den Ouden (2002).

5.1 Participants

SC was a 58-year-old male, who suffered a left-hemisphere stroke 1½ years before testing, resulting in a right hemiparesis and severe aphasia. Testing with the Dutch version of the *Aachen Aphasia Test* (AAT) (Graetz et al., 1992) confirmed the clinical diagnosis of Wernicke's aphasia. In spontaneous speech, SC produced many semantic and literal paraphasias, while his articulation and prosody were judged as 'normal'. Though fluent, with a speech rate of over 90 words per minute, SC used many stereotypes and short, often syntactically incomplete sentences, in which many functional elements were missing. Repetition was moderately impaired, with most problems in longer words and sentences. Writing, naming and comprehension were severely impaired. With respect to the model presented in Figure 13.1, SC is assumed to have a deficit at the lexical level.

GP was an 81-year-old male, who had multiple transient ischemic attacks (TIAs) in the left-hemisphere, resulting in aphasia, 4 months before testing. Testing with the AAT yielded a diagnosis of severe conduction aphasia. Spontaneous speech was fluent (over 90 words per minute), with no indication of an articulatory or prosodic disorder, but with many literal paraphasias, neologisms and *conduites d'approche*. Repetition was severely impaired, mostly on long words and sentences. Written language skills were only moderately impaired and object naming was mildly to moderately impaired, errors being literal paraphasias, neologisms or fairly adequate descriptions. Overall, there was a clear length effect on GP's production errors. At the word level, comprehension was only minimally impaired, but more problems arose at the sentence level. With respect to Figure 13.1, GP seems to have a postlexical deficit, at the level of phonological encoding.

5.2 Materials, Procedure and Scoring

Sixty-two lexical (noun) items were used in three different tasks, viz. naming, repetition and productive phoneme detection. These items had all passed a naming consistency test with unimpaired control speakers, and were split

into nine categories, based on word-form, as shown in Table 13.1. The categories were matched for the mean lexical frequency of their items.

Table 13.1 Experimental item categories

Structure ^a	Stress Pattern	Example	Number of Items
CVCV—	0 1	<i>gitaar</i> (guitar)	7
CVCV—	1 0	<i>tegel</i> (brick)	7
CCVCVC	(variable)	<i>kruimel</i> (crumb)	8
CVC.CVC	0 1	<i>balkon</i> (balcony)	7
CVC.CVC	1 0	<i>panter</i> (panther)	7
CV—.CCV—	(variable)	<i>citroen</i> (lemon)	3
trisyllabic	1 0 0	<i>tekening</i> (drawing)	8
trisyllabic	0 1 0	<i>parade</i> (parade)	8
trisyllabic	0 0 1	<i>sigaret</i> (cigarette)	7

a. **Bold** consonant positions are used in the positional markedness error analysis of phoneme detection task; V— rhymes have variable structures, and may contain empty codas.

The naming task consisted of 62 line drawings of the items. Participants were asked to describe in one word – a noun – what was depicted. In the repetition task, the participants were asked to repeat what was said by the examiner. During this task, the examiner shielded his mouth from the participant's view, in order to minimize the chance that the patient might benefit from lip reading and perform a visual imitation, rather than a lexical repetition. Responses to these tasks were recorded and categorized as 'literal paraphasias' (lemon → *lebon*), 'neologisms' (lemon → *korgewin*), 'semantic paraphasias' (lemon → *banana*), 'irrelevant (verbal) paraphasias' (lemon → *cigarette*), 'no response', 'circumlocutions' (lemon → *very sour*) and 'mixed errors' (lemon → *banala*).

The productive phoneme monitoring task consisted of the same pictures as those used in the naming task, with the addition of graphemic representations of two consonants, one of which would be part of the target item's phonological form (see Figure 13.2). The examiner would sound out the two phonemes, after which participants would have to point to the grapheme representing the target sound, as a forced choice, without articulation. In pseudorandom order, all items were presented twice, with different target phonemes, so that the scores on sounds in different positions in the word could be compared and effects of positional markedness could be assessed. The productive task should be distinguished from the standard use of the phoneme monitoring task in perceptual experiments, where participants monitor an auditorily (or graphemically)

presented string for a target sound (Cutler, 1976). It is more similar to the productive phoneme and syllable monitoring task introduced by Wheeldon and Levelt (1995), in which participants silently translate a presented word into their own language and monitor the resulting string for target sounds. As in the presently discussed experiment, picture-based productive monitoring tasks have also been used by Schiller et al. (2003) and Jansma and Schiller (2004), in the investigation of phonological structure in pre-articulatory word forms.

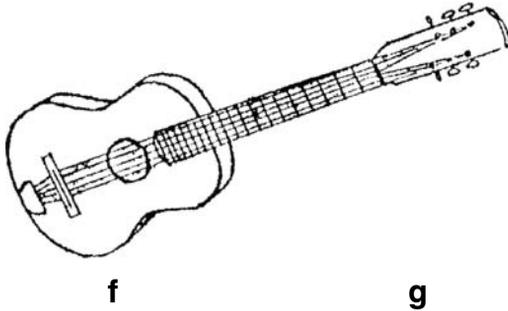


Figure 13.2 Phoneme detection task, target 'gitaar' (guitar) /xita:r/

Items for which the subject intended or produced a semantic paraphasia, irrelevant paraphasia or mixed error, were not taken into account in the analysis of responses on the phoneme detection task. For the measurement of positional effects in the phoneme detection task, the following groupings and comparisons were made of the target structures:

- a. Early vs. late positions
- b. Strong vs. weak positions
 - strong: CCVCVC; CV-.CCV-; CVC.CVC
(all non-sonorant onset heads)
 - weak: CCVCVC; CV-.CCV-; CVC.CVC
(sonorant onset satellites and syllable codas)
- c. Stressed vs. unstressed syllables
- d. Simple vs. complex word structures
 - simple: CVCV-
 - complex: CCVCVC; CVC.CVC; CV-.CCV-
- e. Disyllabic words vs. trisyllabic words.

For the comparison of the forced choice phoneme detection accuracy scores with scores on repetition and naming, a correction for chance was applied, by subtracting the absolute number of incorrect responses from the absolute number of correct responses.

5.3 Results

Figure 13.3 shows the accuracy results for both patients, while Table 13.2 breaks down their phoneme detection scores. For SC, repetition was significantly better than naming and phoneme detection, with no difference between phoneme detection and naming. Most prominent among the errors on the naming task were literal paraphasias (29%), no responses (29%) and circumlocutions (19.4%). On the repetition task, all errors were literal paraphasias. No effects of word-form were noted on SC's errors.

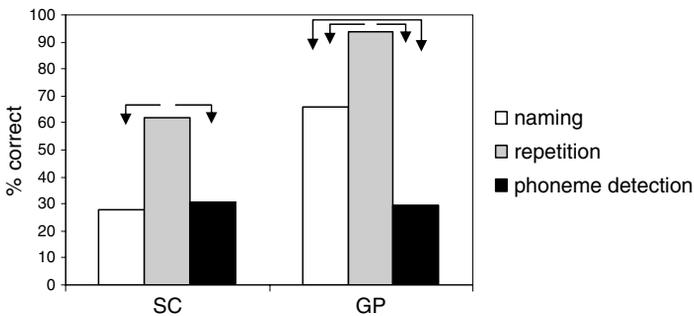


Figure 13.3 Accuracy on naming, repetition and phoneme detection for SC and GP. Phoneme detection scores have been corrected for chance. Arrows indicate significant accuracy differences (χ^2 ; $p < 0.05$)

Table 13.2 Phoneme detection scores, by categorization on structural variables. Given p-values are based on χ^2 tests on raw scores

		SC		GP	
		% correct	p-value	% correct	p-value
Length	early	61.1	0.62	67.9	0.64
	late	69.4		61.8	
Positional markedness	strong	64.3	0.80	71.4	0.09
	weak	61.5		40	
Stress	stressed	66.7	0.90	69.2	0.77
	unstressed	71.4		61.5	
Word complexity	simple	75	0.63	80.8	0.07
	complex	63		56.1	
Syllable number	2 syllables	67.4	0.83	65.7	0.99
	3 syllables	62.1		63.6	

GP performed better on the repetition task than on naming and phoneme detection, while naming was also better than phoneme detection. In naming, literal paraphasias (47%) and semantic paraphasias (28%) were the most prominent error types. The only four errors on the repetition task were literal paraphasias. On the naming task, GP showed an effect of word stress, disyllabic words with stress on the first syllable (12/14) being named more accurately than those with stress on the second syllable (6/14) ($\chi^2=5.600$; $p<0.05$). On the productive phoneme detection task GP also showed a tendency towards an effect of syllabic markedness on the detection accuracy of target phonemes, with phonemes in syllabically strong positions being detected better than those in weak positions, although this did not reach significance ($\chi^2=2.931$; $p=0.087$). There was a similar tendency towards an effect of word structure complexity, with singleton phonemes (in 'simple' words) being detected better than phonemes that were part of tauto- or heterosyllabic consonant clusters ($\chi^2=3.271$; $p=0.071$).

5.4 Discussion and Conclusion

Both SC and GP show literal paraphasias in speech production, and their accuracy scores in the productive phoneme detection task are almost equal, yet their errors are qualitatively different. The phoneme detection task allows investigation of (conscious) access to specific individual segments in the phonological string, bypassing the execution stage of the production process.

SC appears to have a lexical deficit, based on the original diagnosis, the high number of semantic paraphasias and confirmed by the relatively high scores on repetition versus lexical access tasks (naming and phoneme detection). His single-word literal paraphasias did not show effects of phonological markedness, nor did his errors on the phoneme detection task.

GP's deficit seems to be primarily in phonological encoding – when provided with the sound pattern of single words, he can repeat them quite well, but he has more problems retrieving word forms without a phonological cue, and relatively low awareness of the individual phonemes of which lexical items consist, even when he is able to correctly name these same items in whole-word production. On this naming task, GP is helped if the lexical stress falls on the first syllable, possibly because such a pattern allows him to use the first syllable as a peg on which to hang the rest of the word's segments. By contrast to SC, who shows no particular pattern in his phoneme detection errors, GP shows tendencies for positional markedness effects, with more errors on syllabically weak positions, as well as in words which include complex syllable constituents, that is, consonant clusters.

As such, these data suggest that hierarchical syllable structure is not represented at the lexical level, but only comes into play at a postlexical level

of processing, at which individual segments are mapped onto syllable slots. GP's error pattern reveals the amount and type of phonological structure that plays a role at his functional level of deficit, viz. phonological encoding.

6. Conclusion: Formal Phonology and Aphasia

Studies such as those discussed and presented above are examples of how phonological theory may inform the way in which phonological error patterns in aphasia are analysed, and also show that the aphasic data and the psycholinguistic or even neurolinguistic models used to categorize types of aphasia may in turn suggest adaptations to otherwise purely formal phonological models. Such an approach is of course not exclusive to aphasiology, but is valid for psycholinguistics and neurolinguistics in general.

One interesting issue for phonological theory is how to handle the variance in data such as from aphasic speakers. Optimality Theory, for example, is popular for its intuitive handling of the HoP-HoT phenomenon (i.e. 'Heterogeneity-of-Process – Homogeneity-of-Target'; see McCarthy, 2002). However, in aphasia there may be heterogeneity of the target, because markedness is not homogeneous (emerging from deficits at underlying levels of representation, with level-specific well-formedness preferences). There may be no single markedness hierarchy at work in aphasia, even though unmarked features do go before marked when relative to their own 'kind' (Bernhardt and Stemberger, 2007). Any phonological theory that makes a claim to psychological validity has to be able to give a representational, derivational or other account for phonological/structural systematicity observed in the literal paraphasias produced by aphasic speakers with a phonological deficit.

As is clear from the apparent entanglement of error patterns in apraxia of speech and 'deeper' phonological disorders, the study of phonological disorders shows how closely phonology is related to articulation. It is not surprising, then, that many researchers in this field lean more towards functional approaches to phonology, for example, those which are phonetically or cognitively grounded (e.g. Bybee, 2001; Archangeli and Pulleyblank, 1994), and less to viewing the phonological module as a purely abstract standalone module (cf. Code and Ball, 1988; Baker et al., 1999; Ball et al., 2010).

Part III

NEW DIRECTIONS

This page intentionally left blank

14

Laboratory Phonology

Taehong Cho

Chapter Overview

1. The Phonetics-Phonology Interface and the Emergence of Laboratory Phonology	343
2. The Phonetics-Prosody Interface	346
3. Issues in Prosodic Strengthening in Speech Production	352
4. Issues in Prosodic Strengthening in Speech Perception	362
5. A New Challenge	367

1. The Phonetics-Phonology Interface and the Emergence of Laboratory Phonology

What is phonetics and what is phonology? Answering this question is not as simple as it might seem since both disciplines are concerned with speech sounds and sound systems of human language. A common answer to this question would be that phonetics deals with the description of the physical properties of speech sounds that are continuous and gradient in temporal and spatial dimensions, whereas phonology is concerned with abstract aspects of the sound system, describing how the abstract representations of speech units are formed and organized in a categorical way to make contrastive distinctions among them. The distinction between these two subfields of linguistics crystallized in the tradition of the *Sound Pattern of English* (SPE) (Chomsky and Halle 1968). Within this tradition, an extreme view is that the phonetic component is not part of the grammar but ‘physical’ and ‘automatic’, and thus something that can be studied outside the realm of linguistics.

However, the severance between phonetics and phonology has been steadily eroding since the 1980s as researchers began to become aware, thanks to meticulous and scientific methodologies adopted in studying speech sounds, of the importance of scalar and gradient aspects of speech in understanding the

linguistic sound system. In particular, non-contrastive phonetic events, which had traditionally been understood to be beyond the speaker's control (i.e. as low-level automatic physiological phenomena), have been reinterpreted as part of the grammar since they turn out to be either systematically linked with phonological contrasts or governed by language-specific phonetic rules. For example, non-contrastive vowel duration in English is in fact closely tied with the phonological voicing contrast of the postvocalic consonants. Similarly, stop epenthesis between a nasal and a homorganic fricative (e.g. insertion of [t] in *prince*), which was once considered an inevitable automatic effect of producing nasal-fricative sequences, is now understood to be controlled by the speaker and thus driven by a language-specific phonetic rule. For example, Fourakis and Port (1986) have observed that American English speakers insert a stop most of the time while English speakers in South Africa seldom do.

The necessity of language-specific phonetic rules can also be exemplified in the way languages choose phonetic values for the same phonological categories (cf. Keating, 1984; Kingston and Diehl, 1994; Cho and Ladefoged, 1999). For example, in investigating variation and universals in VOT patterns of voiceless stops in 18 languages, Cho and Ladefoged observed that although languages choose one of the three possibilities (unaspirated, aspirated and heavily aspirated) for the degree of aspiration of voiceless stops, they differ not only in choosing VOT values for unaspirated and aspirated stops (e.g. some languages have extremely polarized VOT values for the two phonological categories while others do not), but they also choose VOT values arbitrarily when they have only one voiceless stop category. Based on these observations, Cho and Ladefoged concluded that the phonetic output of a grammar must contain language-specific components, which differ in the target values they assign for the timing between the oral (stop release) gesture and the laryngeal (voicing) gesture for VOT.

Another important finding is that a number of phonological processes which were traditionally assumed to result in complete neutralization are in fact phonetically partial or gradient, showing that phonetic outputs resulting from some phonological processes may not be identical to phonetic outputs that are not phonologically modified. A good example of incomplete neutralization is the case of stop epenthesis mentioned above. Fourakis and Port (1986) show that the inserted [t] in *prince* is phonetically different from the underlying /t/ in *prints*. Another type of incomplete neutralization is found in some assimilatory processes. For example, an electropalatographic (EPG) study by Zsiga (1995) showed that palatalization of English /s/ before a palatal glide /j/ (e.g. /s/ → [ʃ] /__ j, as in *confe[ʃ]ion*) can be gradient when the /s + j/ sequence is created post-lexically across word boundaries (e.g. *confess your*). Similarly, the [p] in *rih[p] berry* (*right berry*), which results from assimilation to /b/, is indeed acoustically different from the underlying /p/ in *ripe berry* (e.g. Gow, 2002).

A number of experimental studies, including the aforementioned ones, have led to a gradual consensus that many, if not all, phonetic events that are expressed by gradient and scalar spatio-temporal dimensions are in fact controlled by the speaker and should be considered as being governed by either the phonological system or by the ‘phonetic grammar’ of the language (e.g. Keating, 1985).

The increasing awareness of language-specific phonetic rules and the close relation between gradient and categorical aspects of speech have thus sparked a fundamental rethinking as to where we should draw the line between phonetics and phonology, or whether a line between them is even necessary (e.g. Keating, 1985, 1990; Cohn, 1998, 2006, 2007; Hayes et al., 2004, among others). Articulatory Phonology is an example of a theoretical framework in which phonetics and phonology are integrated with a set of unified formal mechanisms – that is, phonological contrasts are directly expressed by coordination of articulatory gestures in temporal and spatial dimensions (e.g. Browman and Goldstein, 1992). Another theoretical framework that assumes no boundary between phonetics and phonology is so-called phonetically driven phonology, in which phonological representations are not viewed as categorical but instead as embodying functional phonetic events such as non-contrastive duration, formant transitions and consonant release characteristics (see among others Steriade, 1993b, 1999c; Flemming, 1995; Boersma, 1998; Kirchner, 1998).

Thus far, I have discussed the phonetics-phonology interface at some length since a proper understanding of the sound systems of human language cannot be achieved without understanding their relationship (for this point, see also Hamann, this volume). Indeed, many phoneticians and phonologists have attempted to bridge phonetics and phonology in the past couple of decades, which has led to the inception and consolidation of the interdisciplinary ‘cooperative’ research community called ‘Laboratory Phonology’ (see Cohn, in press, for a thorough overview of the history and the significance of Laboratory Phonology). The Laboratory Phonology community, which started as a small conference in 1987, is now considered to be a particular ‘approach to investigating human sounds and sound systems, taking as foundational the premise that progress will be achieved more successfully through integrated methodologies’ (Cohn, in press). The concerns of Laboratory Phonology are aptly summarized by the three fundamental research questions that Beckman and Kingston (1990: 1) asked at the first Laboratory Phonology conference back in 1987:

First, how, in the twin processes of producing and perceiving speech, do the discrete symbolic or cognitive units of the phonological representation of an utterance map into the continuous psychoacoustic and motoric functions of its phonetic representations? Second, how should the task of

explaining speech patterns be divided between the models of grammatical function that are encoded in phonological representations and the models of physical or sensory function that are encoded in phonetic representations? And third, what sorts of research methods are most likely to provide good models for the two components and for the mapping between them?

Beckman and Kingston (1990: 3) go on to highlight the necessity of a so-called hybrid research methodology in answering these questions successfully, that is, a methodology that requires ‘experimental paradigms that control for details of phonological structure [and] observational techniques that go beyond standard field methods’. Since then, researchers from different disciplines, including phoneticians, phonologists, psycholinguists, cognitive scientists and speech technologists, have been working towards this goal. Laboratory Phonology has now become codified as a truly multi-disciplinary approach bridging the gap between more theoretical and more empirical approaches to the investigation of the sounds and sound systems of natural language (cf. Cohn, *in press*).

2. The Phonetics-Prosody Interface

Among the various research questions and issues that have been actively addressed by the Laboratory Phonology community are those regarding the relationship between phonetics and prosody. In fact, it may not be too much to say that research in Intonational Phonology (e.g. Bruce, 1977; Pierrehumbert, 1980; Pierrehumbert and Beckman, 1988; Ladd, 1996) has been in the vanguard of work attempting to bridge categorical and gradient aspects of speech sounds, and as such stood at the basis of Laboratory Phonology. The main concern of Intonational Phonology was to bring gradient and varying physical F₀ events into a systematically organized pattern by tying them with phonological representations of underlying tones (or tonal targets) through a mapping of categorical tonal representations to continuous F₀ events, that is, the structure of the tune of a given utterance is assumed to be generated by phonetic interpolation between tonal targets.

The fundamentals of Intonational Phonology have thus been adopted in defining ‘intonationally based’ prosodic structure, which consists of structural entities such as prominence-lending locations (e.g. stressed or accented syllables) and prosodic junctures of different levels (see Section 2.1 for further discussion of prosodic structure; and Jun, 1998, for discussion on intonational versus syntactic approaches to defining prosodic structure). How is such phonologically defined prosodic structure manifested in fine-grained phonetic details? And

how are sounds and sound systems (including application of phonetic and phonological rules) constrained and informed by prosodic structure? These questions have played a key role in research on the phonetics-prosody interface. The general consensus in the literature is that sounds and sound patterns of human languages can never be fully understood without understanding how they interact with prosodic structure in the grammatical system of the language (cf. Shattuck-Hufnagel and Turk, 1996). In the remainder of this chapter I will consider the phonetics-prosody interface in more detail. I will start with an introduction to the structural view of prosody. This is followed by an in-depth discussion on the manifestation of prosodic structure in speech production and its role in speech comprehension.

2.1 A Structural View of Prosody

A fundamental property of utterances is that they are prosodic in nature. An utterance is prosodic in that its segmental make-up is superimposed by prosodic (or suprasegmental) features such as F_0 , duration and amplitude; it is also prosodic in that such prosodic features are employed to build a prosodic structure of the utterance being spoken. Prosodic structure can be defined as 'a hierarchically organized structure of phonologically defined constituents and heads' (Beckman, 1996: 19). On this structural view, prosody reflects both constituent- and prominence-based hierarchies.

The constituent-based hierarchy is built with prosodic constituents. In English, these include Syllable, Foot, Prosodic (or Phonological) Word, Intermediate (Phonological) Phrase and Intonational Phrase (cf. Shattuck-Hufnagel and Turk, 1996). The size of prosodic constituents becomes progressively larger with smaller constituents combining to form immediately larger ones in a hierarchical fashion – that is, one or more syllables are grouped into a prosodic word; one or more words combine to form an immediately larger prosodic constituent, the Intermediate Phrase (=ip); and finally one or more Intermediate Phrases combine to form the Intonational Phrase, which is often assumed to be the largest prosodic unit (e.g. Beckman and Pierrehumbert, 1986; but see Nespor and Vogel, 1986). These prosodic constituents are also regarded as prosodic domains in that they often serve as domains for certain intonational patterns and certain phonological rules (cf. Selkirk, 1984b, 1995b; Jun, 1998).

The prominence-based prosodic hierarchy is constituted with increasing degree of prominence from null stress, secondary lexical stress (if it exists), to primary lexical stress and to sentence stress (nuclear pitch accent) (cf. Liberman and Prince, 1977; Beckman, 1986; Hayes, 1989b; Beckman and Edwards, 1994).

This prominence-based hierarchy is closely intertwined with the constituent-based hierarchy in that the lexically stressed syllable serves as the head of the prosodic word, and the syllable with nuclear pitch accent as the head of the Intermediate Phrase (cf. Beckman and Edwards).

Prosodic structure thus serves a dual function in speech production – that is, prosodic boundary marking by which the hierarchical grouping of prosodic constituents is determined, and prominence marking by which relative prominence among prosodic constituents is determined. A single sentence (with the same lexical content and syntactic structure), however, may be produced with different prosodic structures, conditioned by a complex interaction of various factors, for example, what kind of syntactic structure the sentence is built on; what kind of informational structure the utterance conveys in a particular discourse situation; how many syllables or words are available to form one prosodic domain; and how fast the utterance is produced (e.g. Nespor and Vogel, 1986; Pierrehumbert and Hirshberg, 1990; Jun, 1993; Keating and Shattuck-Hufnagel, 2002). Put differently, different prosodic structures of the same sentence may be constructed online in speech production, conveying a great deal of linguistic and extralinguistic information. From the production perspective, prosodic structure is considered an essential element of speech production, modulating phonetic encoding which fine-tunes phonetic outputs so as to signal the prosodic structure (e.g. Keating and Shattuck-Hufnagel, 2002). From the perception perspective, phonetic signatures of a particular prosodic structure are thought to be exploited by listeners in that they help them to process lexical, syntactic and discourse information (e.g. Cho et al., 2007; Carlson et al., 2001; Couper-Kuhlen and Selting, 1996). Thus, the structural view of prosody assumes that prosodic structure is an important grammatical entity in its own right (Beckman, 1996), and therefore a complete picture of speech production and comprehension cannot be obtained without understanding the interplay between phonetics and prosody at various levels of the production and comprehension process.

2.2 Phonetic Manifestations of Prosodic Structure

As an initial step to understand the phonetics-prosody interplay, an increasing number of researchers have endeavoured to explore how prosodic structure may be phonetically manifested. In particular, phonetic manifestations of prosodic structure have been interpreted in terms of ‘prosodic strengthening’, a phenomenon associated with prosodic landmark locations such as prosodic domain edges and prominent syllables. (Here, ‘prosodic strengthening’ is used as a cover term for spatial and/or temporal expansion that is the result of boundary and prominence marking (e.g. Cho, 2005, 2008; Cho and McQueen, 2005)).

Below, I will discuss how prosodic structure can be phonetically expressed in terms of boundary marking and prominence marking.

2.3 Prosodic Strengthening as Boundary Marking

2.3.1 *Domain-Final Prosodic Strengthening (At the Right Edge of Prosodic Domains)*

One of the Most consistent phonetic correlates of prosodic structure is temporal expansion of domain-final segments. The degree of this is closely correlated with the prosodic level or the prosodic boundary strength at the prosodic juncture (e.g. Edwards et al., 1991; Wightman et al., 1992; Gussenhoven and Rietveld, 1992; Berkovits, 1993; Byrd, 2000; Cambier-Langeveld, 2000; Byrd et al., 2006; Cho, 2006). Temporal expansion at the right edge of prosodic domains is usually accompanied by intonational marking, referred to as boundary tones (Beckman and Pierrehumbert, 1986; Pierrehumbert, 1980). Thus, a general consensus is that domain-final articulation is characterized primarily by temporal expansion, which together with boundary tones demarcates prosodic boundaries.

More recently, however, domain-final elements have also been shown to be accompanied by spatial expansion in some cases. For instance, the amount of linguopalatal contact for a domain-final (pre-boundary) vowel (as measured by Electropalatography, EPG) decreases as the boundary level increases (Fougeron and Keating, 1997). A decreased EPG contact indicates more vocalic opening at the end of larger prosodic domains. Subsequent studies with a magnetometer (Electromagnetic Midsagittal Articulography (EMA)) have revealed more evidence for spatial expansion of articulation in domain-final positions, for example, larger C-to-V displacement in French (Tabain, 2003), higher tongue position for /i/ but lower for /a/ in English (Cho, 2005), larger lip aperture for /a/ and /i/ in English (Cho, 2006). Cho (2004) has also shown coarticulatory resistance of domain-final /a, i/ in English, which has been interpreted as another type of articulatory strengthening. These studies thus suggest that domain-final elements may involve spatial expansion, though not as robustly as temporal expansion, the former often being inconsistent across speakers (cf. Byrd et al., 2006) or non-observable (Edwards et al., 1991; Beckman et al., 1992).

2.3.2 *Domain-Initial Prosodic Strengthening (At the Left Edge of Prosodic Domains)*

Another Prosodic landmark location is the other side of the prosodic domain, domain-initial position. This has been considered another locus

for spatio-temporal expansions which mark prosodic boundary strength. Fougeron and Keating (1997) and others (Cho and Keating, 2001; Fougeron, 2001; Keating et al., 2003), for example, used electropalatography (EPG) to measure the degree of linguopalatal contact of domain-initial consonants, and showed that the strength of the consonant articulation, as reflected in degree of oral constriction and seal (closure) duration, increases cumulatively for each higher level in the prosodic hierarchy. Subsequent work has demonstrated similar domain-initial strengthening in various acoustic and articulatory dimensions across languages (Byrd and Saltzman, 1998; Byrd et al., 2000; Lavoie, 2001; Kim, 2003; Cho, 2005, 2006; Cho and McQueen, 2005; Byrd et al., 2006; Kuzla et al., 2007; Cho and Keating, 2009, among others). For example, it has been found cross-linguistically that aspirated stops are produced with longer VOTs in domain-initial than in domain-medial position (for English, see Cole et al., 2007; Cho and Keating, 2009; for Korean, see Jun, 1993, 1995; Cho and Jun, 2000; Cho and Keating, 2001; for Japanese, see Onaka, 2003; Onaka et al., 2003; for Taiwanese, see Hsu and Jun, 1998; Hayashi et al., 1999; for French, see Fougeron, 2001). The longer, that is, temporally expanded, VOTs can be interpreted as a consequence of strengthening of the glottal abduction gesture (Pierrehumbert and Talkin, 1992; cf. Cooper, 1991). In a fiberoptic study, Jun et al. (1998) indeed found larger glottal apertures in AP-initial position than in AP-medial position. (AP is short for Accentual Phrase, which is assumed to be an intermediate prosodic level in Korean; see Jun, 1993, 1995.)

A number of proposals have been advanced to account for the articulatory nature of domain-initial strengthening. Given the close relationship between temporal and spatial expansions associated with domain-initial position, Cho and Keating (2001) proposed an Articulatory Undershoot Hypothesis: in domain-initial position there is enough time to execute the articulatory action, which results in the full attainment of the assumed articulatory target; in domain-medial position, on the other hand, there is some articulatory undershoot due to the insufficient durations associated with these positions. Fougeron (1999) suggested that domain-initial strengthening is ascribable to 'articulatory force' (Straka, 1963), which can be defined as 'the amount of energy necessary to the realization of all the muscular effort involved in the production of a consonant' (Delattre, 1940, translated). In domain-initial position consonants are produced with greater articulatory force, which causes contraction of the muscles involved in the articulation. Based on this assumption, domain-initial strengthening has often been referred to as domain-initial 'articulatory' strengthening (Fougeron and Keating, 1997; Cho and Keating, 2001; Keating et al., 2003). Greater articulatory force in domain-initial position is supposed to apply to the following vowel in such a way as to enhance the aperture contrast between C and V. (See Section 3.1.1 for further discussion

on CV enhancement.) However, it remains as yet unclear whether domain-initial strengthening is strictly confined to initial consonants and, if it is not, how it relates to sonority expansion in those cases where the domain-initial segment is a vowel.

2.4 Prosodic Strengthening as Prominence Marking

Accent-induced phonetic correlates are considered to be another phonetic hallmark of prosodic structure (de Jong, 1991; Beckman et al., 1992; Fowler, 1995; Erickson, 2002; Mooshammer and Fuchs, 2002; Cho, 2006, among others). Prominence marking by accent (sentence stress or nuclear pitch accent) is acoustically manifested with greater F0 movement, longer duration, greater amplitude and unreduced vowel quality (cf. Lehiste, 1970; Beckman, 1986). Articulatorily, it is accompanied by an increased respiratory effort. Based on EMG (Electromyogram) studies, Ladefoged and his colleagues (e.g. Ladefoged, 1967; Ladefoged and Loeb, 2002) showed that syllables with stress, both lexical and sentential, are produced with increased respiratory power due to additional activity of the internal intercostals, which may be closely linked to some characteristics of suprasegmentals under accent. For example, increased subglottal pressure due to heightened respiratory power would result in increased amplitude (see Lehiste, 1970, for discussion).

Prominence is also marked by supralaryngeal articulation, which as Cho (2006) puts it, is ‘simply bigger in all ways – in distance, time, and speed.’ For example, the jaw opening gesture under accent is associated with an increase in duration and displacement, sometimes with a faster movement speed (Fowler, 1995) and sometimes without it (Beckman et al., 1992). The C-to-V lip opening gesture under accent has also been found to be associated with an increase in spatio-temporal expansion along with a faster movement speed (Cho, 2006).

2.5 Boundary Versus Prominence Marking

One of the important questions that has been explored by researchers in the past few decades is whether speakers differentiate between boundary and prominence information in speech production, and if so, how. This question is motivated by the assumption that prosodic strengthening serves a dual function (i.e. boundary marking and prominence marking). If it does, then speakers may feasibly differentiate between the two kinds of strengthening, given that they signal different aspects of prosodic structure. As discussed in the previous sections, phonetic signatures of boundary and prominence markings are

similar in some aspects and different in others. To recap, the most obvious similarity has been found in temporal expansion, in that both boundary and prominence markings are accompanied by lengthening. Both also show some spatial dimension, but differ in that prominence marking is more likely to come with expansion of vocalic displacement whereas boundary marking tends to be characterized primarily by heightened consonantal constriction domain-initially and sometimes expanded vocalic displacement domain-finally. Further, speed of articulatory movement is consistently higher under accent, whereas articulatory movements are not necessarily faster when marking prosodic boundary, neither domain-initially nor finally. These findings, as they are now understood, clearly suggest that the two kinds of prosodic strengthening are indeed differentially manifested in speech production, which supports the view that they are separately encoded in speech planning (cf. Keating, 2006; Cho and Keating, 2009); see Section 3.3 for further discussion on this.

3. Issues in Prosodic Strengthening in Speech Production

Thus far, I have discussed some basic phonetic correlates of prosodic structure which can be broadly characterized as prosodic strengthening as a function of boundary and prominence markings. In the following subsections I will discuss some important issues regarding the production aspects of prosodic strengthening that have been considered in speech prosody literature.

3.1 Prosodic Strengthening and Linguistic Contrast Maximization

From a phonological point of view, a fundamental question is how prosodic strengthening relates to linguistic contrast. An important assumption is that if segments are articulatorily strengthened in prosodic landmark locations, this could result in a linguistically significant heightening of phonetic ‘clarity’, and hence be associated with enhancement of linguistic contrast. Given the observed phonetic differences between boundary and prominence markings, researchers have proposed hypotheses about different types of linguistic contrast maximization, which can be said to fall under the rubrics of *syntagmatic* and *paradigmatic* contrast enhancement (Beckman, 1996; Hsu and Jun, 1998; Cho and Jun, 2000; see Fougeron, 1999, for a review).

3.1.1 Syntagmatic Contrast Enhancement

Boundary marking has often been thought to be syntagmatically, or structurally, motivated, in that it results in enhancement of contrast between neighbouring

segments at prosodic junctures (see Fougeron, 1999, for a review). There is a clear connection between this type of enhancement and domain-initial strengthening. Previous studies suggest that what is strengthened domain-initially is the ‘consonantality’ of the segment. An increased oral constriction degree augments #CV displacement domain-initially and V#C displacement across a prosodic boundary. Longer closure duration and aspiration augment the voicelessness of oral consonants (making them less sonorous), and therefore enhance the contrast with the sonority of neighbouring vowels. Similarly, reduced nasal airflow (Fougeron, 2001) and nasal energy (Cho and Keating, 2009) of /n/ in domain-initial position make this consonant less sonorous and therefore more consonantal. Furthermore, vocalic opening expansion, which is often observed in both pre-boundary and post-boundary positions, can be interpreted as sonority expansion of vowels, which, together with augmented consonantality, contributes to the enhancement of #CV and V#C contrasts.

3.1.2 *Paradigmatic Contrast Enhancement*

Contrary to boundary-induced syntagmatic enhancement, prominence marking is thought to give rise to paradigmatic contrast enhancement, that is, maximization of the phonemic distinction of contrastive sounds. De Jong (1995) refers to this as ‘localized hyperarticulation’. He observed that the English vowel /u/ is produced with a lowered jaw and tongue when accented. This is consistent with an earlier account of prominence marking – the sonority expansion hypothesis – which predicts that the amount of mouth opening, which is related with amplitude, is greater under accent (Beckman et al., 1992, Edwards and Beckman, 1988). However, de Jong also observed that the tongue body is more retracted, enhancing the backness feature of the vowel, which cannot be entirely ascribable to sonority expansion. He therefore proposed that certain accented segments are produced with more extreme articulatory movements in a direction that results in an enhancement of the distinctive features of segments. (See also Cho, 2005, who showed that accent induces fronting of the tongue for English /i/, enhancing the advancement feature.) These, then, are examples of localized hyperarticulation, which is based on Lindblom’s (1990) notion of hyperarticulation, but is assumed to apply not just to extended discourses, but also locally to individual syllables. Thus, it can be used to characterize the extreme articulations in stressed/accented syllables, which ultimately lead to maximization of lexical distinctions. In a similar vein, Fowler (1995) also suggested that, based on the assumption that stress consists of a global increase in production effort (e.g. Öhman, 1966; Lehiste, 1970), accent-induced prosodic strengthening is driven by the prominence maximization principle, reflecting speakers’ ‘global effort’ to increase the perceptual saliency of accented elements for the benefit of listeners.

3.1.3 Language Specificity and Contrast Enhancement

De Jong's localized hyperarticulation hypothesis is based on data from English, a stress-timed language in which lexical stress interplays with sentence stress. However, there are also languages that do not have any word-level prosody (e.g. lexical stress with nuclear pitch accent superimposed on it) and which therefore have a different prominence hierarchy than English has. A question that arises, then, is whether such languages (Korean is an example) differentiate linguistic contrast enhancements in a similar way as English does. This is an open question, which requires further research.

However, recent studies on Korean show some evidence that boundary marking may also be characterized by paradigmatic contrast enhancement (in line with localized hyperarticulation), at least as far as consonantal articulation is concerned. For example, in an acoustic-aerodynamic study Cho and Jun (2000) examined variations in VOT and airflow for domain-initial Korean stops, which display a three-way contrast between lenis, fortis and aspirated (cf. Cho et al., 2002). Cho and Jun found that for lenis and aspirated stops both VOT and the amount of airflow were greater domain-initially than domain-medially, that is, were more consonantal, in line with syntagmatic CV contrast enhancement. However, domain-initial fortis stops, which were produced with reduced VOT and airflow, showed the opposite pattern. The asymmetric domain-initial effects on laryngeal articulation (as reflected in VOT and airflow data) among different stops were interpreted as evidence for enhancement of laryngeal features. Following Lombardi (1991a and b), it was assumed that the three-way stop contrast in Korean is specified in terms of two privative laryngeal features, [spread glottis] and [constricted glottis]: aspirated and fortis stops are specified by [spread glottis] and [constricted glottis], respectively, whereas lenis stops are unspecified for either feature. The increased VOT and airflow for domain-initial aspirated stops were then viewed as enhancement of [spread glottis] whereas the reduced VOT and airflow for domain-initial fortis stops as an enhancement of [constricted glottis]. Cho and Jun further proposed that the unspecified lenis stop was still strengthened, but this time driven by syntagmatic CV contrast, explaining the increased VOT and airflow associated with domain-initial lenis stops. Importantly, the range of VOT and airflow variation for lenis stops rarely overlapped with that of aspirated stops, suggesting that the lenis stops' syntagmatic contrast enhancement was constrained in such a way that it does not blur the paradigmatic contrast between aspirated and lenis stops. (See Hsu and Jun, 1998 for a similar discussion on VOT variation in voiced, unaspirated and aspirated stops in Taiwanese.)

Cho and McQueen (2005) further asked whether domain-initial strengthening is constrained by language-specific phonological factors. They compared domain-initial strengthening in Dutch and English. Both languages have a two-way phonological contrast in stops (voiced versus voiceless), which can

be specified with phonological features [\pm voice] (e.g. Kingston and Diehl, 1994). However, they differ in how the two stops are mapped onto phonetic categories. As Keating (1984) proposed, only three phonetic categories are needed to account for the phonetic realizations of a two-way phonological voicing contrast between stops in the world's languages: voiced, voiceless unaspirated and voiceless aspirated. The Dutch stops map onto the voiced and voiceless unaspirated categories, whereas English stops generally map onto as voiceless unaspirated and voiceless aspirated categories. Crucially, Cho and McQueen found shorter VOTs for Dutch /t/ in domain-initial position, as opposed to longer VOTs in that position in English. Based on the asymmetrical VOT patterns between the two languages, they proposed that prosodically driven phonetic realization is bounded by language-specific constraints on how phonetic features are specified with phonetic content: shortened VOT in Dutch reflects enhancement of [-spread glottis], while lengthened VOT in English reflects enhancement of [+spread glottis]. The shortened VOT in Dutch was also reportedly associated with /t/ in accented syllables, reflecting enhancement of [-spread glottis] as prominence marking under accent, which is again different from the lengthened VOT under accent in English.

Taken together, these results indicate that language-specific phonetic features are enhanced not only as boundary marking but also as prominence marking. Cho and McQueen (2005) proposed that there are cross-linguistic differences in the prosodic modulation of segment realization: the language-specific phonetic component of the grammar (cf. Keating, 1984) modulates the phonetics-prosody interplay. They concluded that prosodic strengthening is not simply a low-level phonetic event but a complex linguistic phenomenon which gives rise not only to enhancement of phonological/phonetic features, but also expresses positional strength, which may license phonological contrasts.

Some phonologists have indeed considered prosodically strong positions as 'privileged' or 'licensing' positions. In these positions, phonological contrasts are most often maintained and segments act as triggers of phonological modification of neighbouring segments (e.g. vowel harmony) but they themselves resist such a modification (e.g. Beckman, 1998; Steriade, 1999c; Barnes, 2002). V-to-V coarticulatory resistance found in prosodically strong locations may also be viewed as a result of contrast maintenance (Cho, 2004). However, as Cho (2005) points out, these approaches do not make clear whether such a positional privilege is phonetically grounded and attributable to the richness of the phonetic cues associated with that position (Steriade, 1999c) or structurally driven and purely attributable to the position itself (Beckman, 1998). Nonetheless, what is clear is that prosodic strengthening is closely linked to maintenance or maximization of phonological contrasts, and is modulated by the language-specific sound system.

3.2 Models of Prosodic Strengthening

In the previous section I discussed how prosodic strengthening can be related to linguistic contrast enhancement. Such enhancements, however, can be achieved only through fine-tuning of articulation of particular segments. In this section I will discuss some possible mechanisms which might further illuminate the nature of prosodic strengthening, with special reference to domain-initial strengthening. An important question is how prosodic structure influences the detailed operation of the articulators involved in strengthening.

3.2.1 Prosodic Strengthening as a Result of Dynamical Parameter Settings in a Mass-Spring Gestural Model

In Articulatory Phonology (e.g. Browman and Goldstein, 1990, 1992), which is based on a mass-spring task dynamic model (e.g. Saltzman and Munhall, 1989; see Hawkins, 1992, for an overview for non-specialists), linguistically significant vocal tract constrictions, dubbed 'gestures,' are viewed as the primitives of phonological representation. In the model, phonological contrasts are maintained in terms of what kind of gestures are involved and how they are temporally coordinated. From the point of view of task dynamics, gestures are described in terms of the behaviour of the abstract 'mass,' which is connected to a 'spring,' and a 'damper' in a critically damped mass-spring system. As Hawkins describes it, it is as if one end of the spring were fixed at the mass, and the other end to the gestural point attractor. As the point attractor moves to a different target location, the spring is stretched, and the mass is pulled towards the target location. The damper makes the mass-spring system critically damped, such that the gesture is generally realized as a one-directional movement towards the target: The mass does not oscillate due to the damping (i.e. it never reaches the target location, and is not pulled back to its original location), but it stays in the target region, continuously and slowly reaching the equilibrium position of the spring (therefore the mass asymptotes towards the equilibrium position).

In the model gestures are specified with a set of dynamical parameter values. Some of the relevant dynamical parameters include target (underlying amplitude), stiffness (or natural frequency) and intergestural timing. Execution of gestures is characterized by articulatory movements, which vary depending on values of these parameters specified for a given gesture. The stiffness parameter reflects the stiffness of the spring, such that the stiffer the string, the faster the articulatory movement; the target parameter determines the target location, such that the larger the target, the more displacement of the articulatory movement; the intergestural timing determines how early or late the following gesture is timed with the preceding gesture, such that the earlier the following gesture comes, the more overlap between the adjacent gestures

and thus the more truncation of the preceding gesture. (For more recent development of Articulatory Phonology, see Goldstein et al., 2006; Goldstein, Pouplier et al., 2007; Goldstein, Chitoran et al., 2007; Saltzman et al., 2008.)

Since its introduction in the late 1980s, researchers working in Articulatory Phonology have attempted to account for prosodically conditioned speech variation in terms of gestural dynamics (e.g. Edwards et al., 1991; de Jong, 1991; Beckman et al., 1992; Harrington et al., 1995; Saltzman, 1995; Byrd and Saltzman, 1998, 2003; Byrd et al., 2000; Byrd, 2000; Byrd et al., 2006; Cho, 2006, 2008), and suggested that prosodically conditioned articulatory variation may be controlled by dynamical parameter settings.

As to accent-induced articulatory variations, some researchers found that articulation under accent is accompanied by an increase in duration and displacement without a concomitant change in peak velocity (see, for example, Edwards et al., 1991; Beckman et al., 1992). As no change in peak velocity in kinematics can be interpreted as a result of no change in stiffness, they suggested that accent-induced kinematic variations are better accounted for by a change in intergestural timing. However, later kinematic studies indicated that while the intergestural timing parameter may be involved with articulatory movements under accent (Harrington et al., 1995; de Jong, 1991), it cannot be the sole controlling parameter as increased movement displacement under accent is often associated with an increase in movement velocity, which can be better accounted for by a change in the target parameter – that is, the larger the target, the faster the movement (cf. de Jong, 1991; Fowler, 1995; Cho, 2006). Recall that Cho (2006) characterizes accent-induced articulatory strengthening as ‘big in all ways’ in that it is accompanied by larger, longer and faster articulatory movements, which suggests that rather than a particular dynamical parameter setting, multiple parameters must be involved in accent-induced articulatory variation.

Turning to boundary-induced strengthening, one question that has been raised is whether the same dynamical account of accent-induced kinematic variations may apply to boundary-induced articulation. Results of some previous kinematic studies (Edwards et al., 1991; Byrd and Saltzman, 1998; Byrd et al., 2000; Byrd, 2000) have converged on the conclusion that the dynamical mechanism governing boundary-induced strengthening is indeed different from that of accent-induced strengthening. A change in the stiffness parameter has been considered as a plausible explanation for boundary-induced strengthening (as opposed to either intergestural timing or target changes, which in some cases may better account for accent-induced strengthening). For example, the lip opening and closing gestures for English bilabial /m/ at edges of prosodic domains showed that movement duration and time-to-peak velocity were highly correlated (Byrd and Saltzman, 1998), suggesting that a local stiffness change may be the source of variation in boundary-adjacent lengthening: the

higher the boundary, the less the gestural stiffness (see also Edwards et al., 1991, for the domain-final jaw closing gesture and Byrd (2000) for the transboundary tongue movement from /a/ to /i/). But again the stiffness account cannot be the complete story since it fails to explain the often observed boundary-induced spatial expansion. Nevertheless, taken together, these studies suggest that speakers must exploit differential dynamical mechanisms, which distinctively govern prominence versus boundary markings. However, it is still unclear whether and exactly how these are controlled differentially by particular sets of dynamical parameters.

3.2.2 π -Gesture as a Device Modulating Boundary-Adjacent Articulation

As previous studies have not been entirely successful in pinpointing the exact dynamical parameter setting underlying boundary-adjacent articulation, Byrd and her colleagues (Saltzman, 1995; Byrd et al., 2000; Byrd, 2000, 2006; Byrd and Saltzman, 2003; Byrd et al., 2006) have proposed that boundary-induced articulation can be better understood as a result of the influence of a so-called π -gesture that is governed by prosodic constituency in the task dynamics model. The π -gesture is an abstract and non-tract variable 'prosodic' gesture, which determines articulatory movement speed, modulating the rate of the clock that controls articulatory activation of constriction gestures. (A 'non-tract variable' is a gesture that is not actually realized in terms of vocal tract constrictions.) As the clock, controlled by the π -gesture, is slowed down at a prosodic juncture, the articulatory movement at the juncture becomes slower, and possibly spatially larger. In temporal domains the π -gesture is anchored at a prosodic boundary, such that its clock-slowing effect is stronger at the juncture, dwindling farther from the edge. In line with this, some recent kinematic studies have shown that boundary-induced lengthening is most robustly manifested in the 'transboundary' articulatory movements (movements that start domain-finally and end domain-initially spanning the intervening juncture, as in /i/-to-/a/ movement in /i#Ca/ context), which are assumed to be under a direct influence of the π -gesture (see also Byrd et al., 2006 and Cho, 2006, 2008, for English data; Tabain, 2003 and Tabain and Perrier, 2005, for French data).

However, Cho (2008) observes that the π -gesture approach raises a number of issues. One concerns the question exactly how far the boundary-adjacent temporal influence can be extended, in particular with respect to the right of the prosodic boundary. Put differently, is the boundary effect on duration strictly local to domain-initial position? Previous studies have in fact shown mixed effects on domain-initial temporal expansion, suggesting that domain-initial strengthening is not strictly local to 'initial' consonantal articulation. On the one hand, quite a few studies (Fougeron and Keating, 1997; Cho and Keating, 2001;

Keating et al., 2003; Cho and McQueen, 2005; Cho and Keating, 2007, 2009) demonstrated that boundary effects are mainly local to C in domain-initial CV. On the other hand, some kinematic studies have indicated that articulatory movement lengthening effects can be pervasive into the vocalic articulation after the initial consonant (Byrd, 2000; Cho, 2006, 2008; Byrd et al., 2006). For example, Byrd et al. found lengthening of the articulatory opening movement from C to V after a boundary (though this was not observed for all speakers) and suggested that articulatory lengthening effects are roughly equivalent in both pre-boundary (domain-final) and post-boundary (domain-initial) articulations, especially as far as articulations immediately adjacent to the juncture are concerned. (Note, however, that post-boundary compensatory shortening was found, especially for consonantal gestures in onsets, in the second and the third syllables; cf. Krivokapić, 2007.) Cho (2006, 2008) also showed lengthening of vocalic gestures in domain-initial #CV lip opening movement and V-to-V tongue movement across a boundary (V#CV). These mixed results have been interpreted as suggesting that domain-initial effects on duration are not strictly local to the edge but may be gradient in nature, as a function of the distance from the boundary (Cho, 2008; Cho and Keating, 2009).

This gradience is indeed what the π -gesture predicts, that is, its influence wanes as it gets farther away from the boundary. Nevertheless, the mixed results do prompt further questions: exactly how far can the effects of the π -gesture be extended around prosodic junctures, and what factors influence the determination of its scope? Although answering these questions requires substantial follow-up studies, a possible determining factor is the language-specific prosodic system. Barnes (2001, 2002) suggested that, in English at least, the vowel in CV syllables is not subject to domain-initial acoustic lengthening because vowel duration in this language is a major cue for lexical stress. (See also Keating et al., 2003 for discussion of how languages with different prosodic systems may show differential domain-initial strengthening effects.)

Another topic of debate concerns the question of how the π -gesture captures possible spatial expansion at prosodic boundaries. While a body of experimental work has characterized the nature of boundary-adjacent articulations in primarily temporal dimension, other studies (as they are now understood) have shown that boundary-induced spatial expansion affects not only initial consonants, but also often the following vocalic articulation. A simulation of a clock-slowness of the π -gesture (Byrd and Saltzman, 2003) showed that clock-slowness can reduce articulatory overlap between domain-initial consonantal gesture and the following vocalic gesture. The resulting non-truncation of the articulatory target of the consonant gesture may explain consonantal strengthening effects (which is reminiscent of the articulatory undershoot account; cf. Cho and Keating, 2001). However, it remains as yet unclear how

this theory can explain vocalic expansion after the consonant; for this, further development of the model is needed.

As has been discussed so far, the π -gesture theory provides a possible way to unify dynamical accounts for symbolic prosodic and segmental units by relating them to the abstract π -gesture and the tract-variable articulatory gestures, respectively. However, further research is required to account for the gradient nature of both temporal and spatial variation, in a way that is both cognitively plausible and computationally implementable.

3.2.3 Bonding Strength

A possible alternative account that is available in a mass-spring dynamical system involves the notion of 'bonding strength', that is, the degree of cohesion among gestures (see Cho, 2001, based on the work of Browman and Goldstein, 1992). An important assumption of this account is that every gestural phase relation is associated with a differential degree of bonding strength. The original motivation for this was to take into account the variability in context-sensitive phasing relationships between consonantal gestures. Greater bonding strength is assumed to give rise to stronger gestural cohesion and more co-articulation between adjacent gestures. Given that this could be extended more generally to other phasing relationships, Cho et al. (2007) discussed the possibility that some types of boundary-induced phonetic variation can be explained on the assumption that bonding strength is inversely correlated with prosodic boundary strength (or the size of the prosodic domain). For example, cross-boundary V-to-V coarticulatory resistance (Cho, 2004) could be due to reduced bonding strength between vocalic gestures across a prosodic boundary. Similarly, reduced bonding strength would result in less coarticulation between a domain-initial consonant and its neighbouring gestures, leading to a greater duration of the former. If, as I propose, bonding strength is inversely co-indexed with boundary strength, then this information could be fed to the phonetic implementation stage, where prosodically driven fine-tuning of articulation may occur.

3.2.4 The Window Model

Prosodically conditioned phonetic variation can also be understood through Keating's (1990) window model (see also Byrd, 1996; Cho, 2004; Cohn, 1990; Docherty, 1992; Keating, 1996; see, for example, Guenther, 1995; Guenther et al., 1998, for an independent development of a window model). In the window model, articulatory movement targets are assumed to vary within specified ranges (i.e. 'windows') as opposed to fixed values. The model has been proposed to capture boundary-induced phonetic variation in such a way that the range of articulations adjacent to a prosodic boundary can be expanded or

shrunk (Keating and Shattuck-Hufnagel, 2002). An alternative formulation, however, was proposed by Cho (2004): the possible range of variation of a given articulation remains fixed, but prosodic factors such as boundaries could specify an operating target region within the fixed window. For example, the target for an articulator could be modulated according to the prosodic structure generated on line, such that the same gesture is specified for a more extreme and narrower operating range in domain-initial than in domain-medial articulation. Cho (2004) suggested that such modulation of the window range can also be applied to prominence-induced articulatory variation. The Window Model thus captures prosodically driven phonetic variation in a linguistically intuitive way: the phonological category of a segment is expressed by an invariable window size, but the prosodically driven variability is allowed within that window. It still remains to be seen, however, how this interpretation of windows can be further developed in a computationally implementable way.

3.2.5 *An Exemplar-Based Model*

A relatively recent type of production model is the exemplar-based approach (e.g. Pierrehumbert, 2001, 2002, 2003). The main tenet of the model is that speech perception involves the storage of exemplars of specific speech events in a multidimensional phonetic space, and that emerging phonetic categories form 'clouds' in different regions of this space. Each member of a cloud is a remembered instance of a given category (or label), such that each category is associated with a frequency-weighted distribution of phonetic events. Crucially, perceptually cumulated exemplars are used in speech production. When a phonetic category label is chosen, the motor commands for that label are executed based on a random sampling from the distribution of exemplars associated with that label. As Cho et al. (2007) suggest, the exemplar clouds could be co-indexed by prosodic domain. For example, the cloud for a consonant could be subdivided into clouds according to the level of prosodic domain (e.g. Intonational Phrase, Intermediate Phrase, Prosodic Word), so that once the category label and its prosodic position have been selected in production, a candidate exemplar could be selected at random from the prosodically appropriate sub-distribution of exemplars associated with that label (though Cho et al., also discuss a number of potential drawbacks of an exemplar-based approach).

3.3 The Prosody Generator in Speech Production

The mechanisms discussed in the previous section have some implications for psycholinguistics theory of speech production. In the theory of production

developed by Levelt and colleagues (Levelt, 1989; Levelt et al., 1999), the influence of the prosodic structure of an utterance on articulatory planning occurs after lexical access, operated by the 'Prosody Generator' device. (See Keating and Shattuck-Hufnagel, 2002, for further discussion on where in the speech planning process the prosodic structure of an utterance should be built.) This device receives abstract phonological input about words along with information about their prosodic structure, and adjusts phonological specifications as a function of the context – a process called 'prosodification', which occurs before the phonetic encoding stage. (Phonetic encoding involves the specification of phonological representations on the phonetic surface, while phonological encoding refers to the phonological specification of lexical representations or 'lemmas'.) The Prosody Generator determines, for example, resyllabification and stress shift. However, in the current model it does not take into account fine-grained phonetic details that correlate with prosodic structure, such as domain-initial strengthening. The output of prosodification is fed into the phonetic encoding stage, where 'preprogrammed' gestural scores (in the sense of Articulatory Phonology) are retrieved from the 'mental syllabary'. Crucially, the model assumes that the retrieval of the gestural scores is the final step in speech planning, after which motor execution takes place as an automatic and biomechanic phonetic process. What has not been considered in the model is how fine-tuning of these 'preprogrammed' gestural scores occurs as a function of prosodic structure. Cho et al. (2007) proposed that such fine-tuning could still be done by the Prosody Generator, but that it must occur at the phonetic encoding stage, given that it must reflect prosodically driven subphonemic variation of speech. (See also Keating, 2006, for discussion of this issue.) To do this, the Prosody Generator could be harnessed with some of the mechanisms discussed in the previous section. Given that the notion of gestural scores adopted by the model is based on a mass-spring gestural dynamical system, it may be theoretically coherent for the Prosody Generator to be further developed by incorporating the mechanisms of the π -gesture model.

4. Issues in Prosodic Strengthening in Speech Perception

As we now understand it, prosodic strengthening is not simply a physical articulatory event beyond the domain of linguistic control, but is realized in a linguistically meaningful way, conveying information about boundaries and prominence. A question that naturally arises, then, is in what ways such structural information is useful in speech comprehension. In this section I will discuss some possible roles of prosody in speech comprehension,

especially in terms of lexical processing, and propose that prosodic information is processed in parallel with segmental information, by a special device called the 'Prosody Analyser'.

4.1 Phonetic Manifestations of Prosodic Structure in Speech Comprehension

Some of the earlier studies on roles of prosody in spoken word recognition have focused on how word-level prosodic information about lexical stress is exploited in lexical segmentation. For example, Cutler and her colleagues (Cutler and Norris, 1988; Cutler and Butterfield, 1992) showed that English listeners tend to use the frequent strong-weak (trochaic) stress pattern to detect the beginning of a word, which of course is cued by suprasegmental features such as F0, duration and amplitude (Lehiste, 1970). Fragment priming experiments in Spanish (Soto et al., 2001) also suggested that when the stress of the spoken fragment is matched with the stress of the visual target, recognition of the target word is facilitated, but inhibition occurs when stress is mismatched. Other studies have examined effects of cues to prosodic word boundaries in lexical access (e.g. Gow and Gordon, 1995; Davis et al., 2002; Salverda et al., 2003). One particular line of enquiry has examined the case of shorter words embedded in longer words (e.g. *ham* embedded in *hamster*). In eye-tracking experiments with cross-spliced Dutch materials, for example, where participants were instructed to click on the object picture (e.g. *ham* or *hamster*) mentioned in the auditory sentence, Salverda et al. showed more transitory fixations to pictures of monosyllabic words (e.g. *ham*) when the first syllable of the target word (e.g. *hamster*) was cross-spliced from the original monosyllabic word (*ham*) than when it was from the longer word (*hamster*). They then argued that the acoustic durational cues were a consequence of a difference in prosodic structure, and that they modulated lexical activation.

The studies discussed above demonstrated word-level prosodic effects on spoken word recognition. Our understanding of how phrase-level prosodic information is used in spoken word recognition is still limited. However, more recently research in lexical processing has begun to investigate possible roles of prosodic information in higher-level prosodic structure (e.g. Christophe et al., 2004; Kim, 2004; Cho et al., 2007; Shukla et al., 2007; Welby, 2007; Kim and Cho, 2009; Warner et al., in press). For instance, in word monitoring and phoneme detection tasks in French, Christophe et al. showed that lexical access for monosyllabic words (e.g. *chat* 'cat') in two-word sequences (e.g. *chat grincheux* 'cat grumpy') was faster when Phonological Phrase (an intermediate level in the prosodic hierarchy) boundaries intervene than when the sequences

occur within such a phrase. They proposed that listeners terminate pending lexical searches when they encounter Phonological Phrase boundaries, and therefore that lexical access operates within the domain of the Phonological Phrase. However, Christophe et al.'s study did not examine exactly what kind of phonetic information about prosodic structure was used in lexical processing, though they speculated that phrase-final lengthening would be a major cue.

In an effort to explore how specific prosodic cues to high-level prosodic structure may be used, Kim (2004) employed an artificial language learning paradigm (for this, see Saffran et al. 1996). Kim created a small artificial lexicon and tested how novel words were learned by Korean listeners. In the learning phase, listeners were exposed to continuous sequences of the novel words for a certain period of time, and in the test phase they had to identify whether a given string of sounds is a likely word that might occur in the artificial language that they had been exposed to. Crucially, in creating spoken novel words, various prosodic cues were manipulated, such as duration (word-finally, in line with phrase-final lengthening), amplitude (word-initially, in line with domain-initial strengthening), and F₀ (rising word-finally, in line with the Korean intonational structure, where an Accentual Phrase ends with a rising tone). When novel words contained these cues in conformity with the specific prosodic characteristics of Korean, learning the novel words was improved as compared to the control case, where no prosodic cues were available. In particular, the presence of AP-final durational cue was found to improve learning the most. On the other hand, when a rising tone, an AP-final intonational marker in Korean, was aligned with word-initial position, learning performance became poorer than in the control case. Kim's study therefore suggested that each of prosodic cues to high-level prosodic structure contributes to lexical segmentation, but only if the cue conforms to the language-specific prosodic structure (cf. See Tyler and Cutler, 2009, for a cross-linguistic study where phrase-final lengthening served as a cross-linguistic cue).

Subsequently, in word-spotting experiments in Korean, Kim and Cho (2009) further tested how such multiple prosodic cues to high-level prosodic structure contribute cumulatively to lexical segmentation. They found that when both phrase-final duration and intonational cues are available, listeners did not necessarily benefit from them both. Instead, the results showed partial cumulative effects: the addition of phrase-final durational cues was most efficiently exploited when intonational information was not optimal (i.e. with infrequent intonational patterns that do not conform to the intonational phonology of Korean; cf. Jun, 1993, 1995, 2000). Taken together, these studies therefore demonstrate that suprasegmental cues which mark prosodic boundaries are

indeed exploited by listeners in lexical segmentation, but the way they are used is constrained by language-specific prosodic structure.

Let us now consider some possible roles of domain-initial strengthening whose characteristics are not entirely suprasegmental, that is, those which are accompanied by extreme articulation. Domain-initial position has often been considered to be an informationally rich locus in speech processing (see Gow et al., 1996). A basic assumption has been that the speaker signals prosodic structure via articulatory domain-initial strengthening, and that the listener makes use of its acoustic consequences in comprehension (cf. Cho et al., 2007). As Keating (2006) observes, domain-initial strengthening is presumably motivated by the fact that initial segments are less determined by prior context, and the contextual information gap is compensated for by domain-initial strengthening. Fougeron and Keating (1997) also speculated that since domain-initial strengthening entails increased articulatory contrast between segments straddling a prosodic boundary, this contrast could contribute to marking that boundary, thereby helping listeners to parse the continuous incoming speech signal into words and thus facilitating lexical segmentation.

Cho et al. (2007) moved beyond speculations, and directly tested the role of domain-initial strengthening in lexical segmentation. In cross-modal identity-priming experiments, listeners made lexical decisions to visual targets (e.g. *mill*) as they heard the sentence with two-word sequences containing lexical ambiguities (e.g. *mill company*, with the competitor *milk*). The two-word sequences (e.g. *mill # company*) contained Intonational Phrase or Prosodic Word boundaries (IP-boundary context, for example, *To learn about wood products, they visited a MILL COMPANY in Alabama last summer* versus Word-boundary context, for example, *When I was thinking about buying a coffee MILL, COMPANY names were the most important things I considered*). The second word's onset (e.g. [kɑ] in *company*) was then spliced from another token of the sequence in IP-initial position (strengthening condition) or Wd-initial position (non-strengthening condition). When related targets were identical to pre-boundary words (e.g. *mill*) in Word-boundary context, a stronger priming effect was observed in the strengthening condition (when the post-boundary onsets were spliced from IP-initial than from Wd-initial position). The strengthened [k] was viewed to provide a better match to the onset of *company* than the coda of *milk*, its competitor in the sequence *mill company*. Domain-initial strengthening present in the onset of the following word was, therefore, interpreted as serving as a cue to lexical segmentation via resolving lexical ambiguity. A general conclusion is that phonetic correlates of domain-initial strengthening are used as acoustic cues in the segmentation of continuous speech, and, more broadly, that speakers signal prosodic structure in systematic, fine-grained phonetic detail and that listeners make use of it in speech comprehension.

4.2 The Prosody Analyser in Speech Comprehension

Successful spoken word recognition necessitates successful lexical segmentation of continuous speech input that is, finding the word boundaries as intended by the speaker. A well-known approach to lexical segmentation is to consider the process as a consequence of lexical competition (e.g. Marslen-Wilson and Welsh, 1978; McClelland and Elman, 1986; Norris, 1994). In lexical competition, a set of competitors whose acoustic beginnings are consistent with the speech input is initially activated. Competitors are then inhibited as they become mismatched with the input as it unfolds over time. Eventually, a single candidate remains as the winner in the competition. This also ends the search for the word boundary, and hence lexical segmentation. The fact that fine-grained phonetic (subphonemic) details influence lexical segmentation (including prosodically driven phonetic variation) has challenged traditional models of lexical competition which rely on phonemic representations only (cf. Gow and Gordon, 1995; Norris et al., 1997; Gow, 2002).

To account for how the fine-grained phonetic information about prosodic structure influences lexical competition, Cho et al. (2007) have proposed the 'Prosody Analyser' account (see also Salverda et al., 2003), as a mirror image of the Prosody Generator account in speech production (see Section 3.3). Their hypothesis was that the Prosody Analyser uses information which specifies suprasegmental (and possibly segmental) aspects of the speech signal to compute the prosodic structure of the current utterance. Crucially, the Prosody Analyser is not completely separated from the process of segmental analysis (which is necessary to retrieve phonemic representations), but representations of prosodic structure are taken to be extracted in parallel to segmental representations. On the one hand, the segmental analysis determines what words are considered (i.e. the content) in the current input, possibly in terms of phonemic representations in the Shortlist model (Norris, 1994) and the TRACE model (McClelland and Elman, 1986), or possibly in terms of fine-grained phonetic details as in the Feature Parsing model (Gow, 2002). On the other hand, the Prosody Analyser guides the process which determines where words are likely to begin and end, for which it uses the location information of boundaries to modulate the lexical competition process. Potential lexical boundaries, computed by their segmental match with the speech input, are then checked against prosodic boundaries computed by the Prosody Analyser, so that the alignment between the two hypothetical boundaries plays an important role in determining the 'winner' in the lexical competition. The Prosody Analyser account is therefore in line with the assumption that segmental and prosodic information have different roles to play in the word recognition process. See Cho et al. (2007) for further discussion on how an episodic account of word recognition (e.g. Goldinger, 1996, 1998; Johnson, 1997) could make use of fine-grained

phonetic details in lexical processing and on possible advantages of the Prosody Analyser account over the episodic account.

5. A New Challenge

This chapter has considered the phonetics-prosody interface in the realm of Laboratory Phonology, focusing on how prosodic structure is manifested in fine-grained phonetic details in speech production and what roles they may play in speech comprehension. We have seen ample evidence that prosodically driven fine-grained phonetic details do not arise simply as low-level physical phenomena, but instead play linguistically significant roles, serving dual functions of boundary marking and prominence marking. We have also learned that phonetic manifestations of prosodic structure are closely linked to contrast enhancement, both paradigmatically (i.e. phonemic) and syntagmatically (i.e. structurally), and that the way linguistic contrasts are enhanced is modulated by language-specific phonological systems. Moreover, we have seen that phonetic consequences of prosodic strengthening are indeed used in speech comprehension, especially in terms of lexical segmentation and resolution of temporal lexical ambiguities. It was proposed that prosodic structure should be computed (e.g. by the Prosody Analyser) in parallel with segmental analysis, and that the match between computations of prosodic structure and lexical content facilitates lexical processing.

Over the past two decades, we have clearly obtained better insights into the nature of sound systems in which phonetics and prosody are intertwined in a linguistically systematic way. Understanding the interplay between phonetics and prosody must now be seen as an essential prerequisite for understanding the whole linguistic communication system of a given language. We are now left with a new challenge: how do we reinforce our knowledge and extend it to build models that capture the complex interplay of phonetics and prosody, in a way that is both descriptively and explanatorily adequate?

In building such a model, possible mechanisms underlying phonetic manifestations of prosodic structure may include the π -gesture model, Bonding Strength, the Window Model and Exemplar-based approaches, which can be combined with a psycholinguistic production model such as Levelt's, about which I proposed that prosodically driven fine-tuning of articulation must operate at the level of phonetic encoding, possibly by the Prosody Generator. The challenge that we are facing is how to put these insights together and integrate them into a more general model of speech production. As Beckman (1996) puts it, prosody is now considered to be a grammatical entity in its own right. In building such an integrated model, the most important issue will be where in the general architecture of grammar prosody needs to be placed

and how the prosody component interacts with other components in the grammar. All in all, since phonetic manifestations of prosodic structure convey information about various strata of speech production (from low-level phonetic realizations to the discourse structure), and since prosodic information is indeed exploited in speech comprehension, the Laboratory Phonology community needs to develop a more full-fledged model. This model should provide an adequate explanation of the interactions between prosody and other linguistic components in order to achieve a better understanding of global linguistic functions of speech prosody in the architecture of grammar.

15

Usage-Based Phonology

Daniel Silverman

Chapter Overview

1. Introduction	369
2. Precursors to Modern Approaches	370
3. Modern Currents in Usage-Based Phonology	380
4. Conclusion	392

1. Introduction

Phonologists must contend with two incontrovertible facts:

1. Phonological systems consist of *discrete* psychological categories
2. Phonological categories emerge from *variable* speech tokens

An approach to phonology may be characterized as ‘usage-based’ to the extent that it investigates the nature of – and formulates compelling hypotheses about the interaction of – these two aspects of phonological structure. This search for explanation in phonology is usually undertaken in one or both of two arenas: (i) the psychological and physical world of the individual, in the form of laboratory investigations and (ii) the social world of language transmission, in the form of quantitative field studies that, broadly construed, include corpus studies, frequency studies, dialectal variation and the intimately related area of sound change. Indeed, two of the modern progenitors of usage-based phonology have established their research niches accordingly. Ohala’s research programme focuses on using the laboratory as a quasi-time machine,

sometimes inducing physical and perceptual conditions that may reflect slow-going natural sound changes in 'speeded-up' form. It has spawned the burgeoning field of experimental phonology. Labov's research programme involves the study of real-world sound change in the form of longitudinal and dialectal quantitative analyses of speech. It has spawned the field of quantitative sociolinguistics.

This chapter is divided into two sections. The first section consists of a survey of early (nineteenth- to mid-twentieth century) observations that may be interpreted as prefiguring the modern, quantitative research programmes pioneered by Ohala and Labov. Among the most important authors to be discussed are Kruszewski, Baudouin de Courtenay and Martinet. The second section outlines the research programmes of Ohala and Labov, and further discusses modern research issues in usage-based phonology that have been directly or indirectly influenced by these two major contributors to phonological theory.

2. Precursors to Modern Approaches

2.1 The Kazan School

The Kazan School of linguistics, based in Kazan Russia and headed by Baudouin de Courtenay and Kruszewski, may be seen as the historic reflex of several strains of linguistic theorizing that have come to prominence in more recent times. As will become clear, the influence on these early thinkers (and, indeed, on contemporary scholars as well) of Darwin's theory of evolution by means of natural selection cannot be overestimated. Kruszewski and Baudouin de Courtenay both reference Darwin in their writings, but even if they had not, the mark of the Darwinian revolution is writ large in their theories of linguistic sound structure, with their emphasis on slow-going diachronic pressures that may shape and re-shape the linguistic system, due to specific patterns of use and disuse.

In his master's thesis of 1881, Kruszewski outlines his proposals on the origins and properties of sound alternations. He divides sound alternations into three types, each with its own constellation of phonetic and functional properties.

The first type of sound alternation consists of regular phonological alternations admitting of no exceptions. For example, as exemplified in (1), 'When Modern German *s* is followed by a vowel, and, at the same time, preceded by a sonant or by zero, it necessarily becomes *z*. This is exceptionless' (1881: 11). (The examples in (1) do not, in fact, exemplify alternation. It is curious that Kruszewski chose these forms when alternating forms are readily available.)

(1) zeele	haus
unzer	ɔks
gewezen	ist

Another example, from Russian, involves the automatic palatalization preceding *e* and *i*, for example, *t* → *tʲ*.

(2) sveta	'of the world'	sve ^{tʲ} e	'in the world'
		tʲeksas	'Texas'

Kruszewski claims that such patterns are fully explicable in articulatory terms. In such alternations, (i) the causes can be immediately identified, (ii) the alternation is exceptionless and (iii) the alternating sounds are usually articulatorily similar.

Regarding the origins of such changes, Kruszewski's remarks in 1883 make it clear that he sees language structure as intimately connected with language use, prefiguring subsequent theorizing in, more generally, Ohalaian and Labovian approaches to sound change, and more specifically for current purposes, listener-based exemplar modelling and probability matching (discussed in detail in Section 2):

In the course of time, the sounds of a language undergo changes. The spontaneous changes of a sound depend on the gradual change of its articulation. We can pronounce a sound only when our memory retains an imprint of its articulation for us. If all our articulations of a given sound were reflected in this imprint in equal measure, and if the imprint represented an average of all these articulations, we, with this guidance, would always perform the articulation in question approximately the same way. But the most recent (in time) articulations, together with their fortuitous deviations, are retained by the memory far more forcefully than the earlier ones. Thus, negligible deviations acquire the capacity to grow progressively greater . . . (1883[1995]: 85)

As an example, let us take the sound *k*^{*l*}. Let us imagine its domain as a line; on one end the sound *k*^{*l*}, will have the slightest palatal nuance (*k*₁^{*l*}), and on the other end a very significant palatal nuance (*k*_{*n*}^{*l*}): the intermediate points (*k*₂^{*l*}, *k*₃^{*l*}, *k*₄^{*l*}) will designate the articulation between *k*₁^{*l*} and *k*_{*n*}^{*l*} . . .

$$k_1^l \dots k_2^l \dots k_3^l \dots k_4^l \dots k_5^l \dots k_6^l \dots k_7^l \dots k_n^l$$

Let us imagine that, while pronouncing k^j , we have articulated k^j_2 . Which articulation will we perform the next time we pronounce the same sound? Each of the articulations is directed by the unconscious memory of similar articulations performed on previous occasions; therefore, we can perform the same articulation of k^j_2 . However, our memory retains only an approximate picture of the previous articulation, and our organs perform only approximately the same operation which we make them perform. Therefore, it is much more likely that the next time we will perform the articulation not of k^j_2 , but of one of its neighboring articulations k^j_1 or k^j_3 . Let us assume that we have performed the articulation of k^j_3 . Which articulation will we perform the third time? Our characteristic, unconscious memory of the articulation of sound k^j should be a complex recollection of all articulations of k^j which we have performed. But not all of these articulations are arranged equally in the memory. For this reason, after performing the articulation of k^j_3 , the chances of performing k^j_4 are much greater than they are for k^j_1 , etc. (1883[1995]: 65–66)

Kruszewski's second category of sound alternation is exemplified by German **s-r: war-gewezen**. Such patterns (i) have many exceptions, and thus cannot be stated in purely phonological terms, (ii) possess causes and conditions that require a paleophonetic and etymological investigation, (iii) are typically associated with only certain morphological concatenations (or concatenation classes), but not necessarily exhaustively, (iv) often have triggers that are unclear, in the sense that the (natural) cause of the alternation has been obscured by historical changes and (v) involve phonetically dissimilar alternants.

Kruszewski's third category consists of morphologically conditioned alternations. He again employs examples from German and Russian. First, consider German umlaut.

(3) German umlaut

h <u>au</u> s	h <u>æu</u> z-e	h <u>æu</u> s-lain
r <u>ad</u>	r <u>æd</u> -e	r <u>æt</u> -lain
l <u>ox</u>	l <u>øx</u> -e	l <u>øx</u> -lain
b <u>ux</u>	b <u>yx</u> -e	b <u>yx</u> -lain

Such patterns (i) require a paleophonetic investigation, (ii) are productive (apply to nonce forms), (iii) are exceptionless within the paradigm in which they are present and (iv), are morphologically conditioned.

Consider also the k - tʃ alternation in Russian.

(4) Russian k - tʃ alternation

pr <u>ə</u> rok	'to profit'	pr <u>ə</u> roftit	'prophesy'
-----------------	-------------	--------------------	------------

All **k**-final nouns have corresponding verbs with **ʃ** as the stem-final consonant. As with German umlaut, this alternation is limited to particular grammatical category changes, and thus serves a morphological function.

While 'all of the phenomena which we have been discussing result from physical processes called combinatory and spontaneous sound change, and from unconscious psychical processes . . .' (Kruszewski, 1881: 19), the grammatical uses towards which particular alternations are put have consequences for their diachronic trajectory. For example, Type-1 alternations involve exceptionless physically-based correlations (**s** : **s₁**). Over time, **s₁** may now become another sound, **z**, and a new correlation **s** : **z** is introduced. Such **s** : **z** patterns admit exceptions, since the alternation is not causal or automatic. A Type-1 pattern may thus diachronically evolve into a Type-2 pattern. Type-2 patterns, in turn, may be further subject to 'psychical' pressures such that (sub-)regularity is re-introduced. A pattern might level such that it limits itself to particular paradigms, and morphological conditioning becomes possible. This is Type-3.

An example of the evolution of changes in alternation types comes from a case of paradigm levelling in Eastern Slavic. Note that an automatic, phonetically explicable pattern of palatalization has become removed from its phonetic origins in Russian, such that the two sounds in question no longer bear a 'physical' or 'psychical' relationship to one another. Subsequently, sub-regularity is reintroduced through paradigm levelling: the alternation levels towards different values in a Russian dialect and in Ukrainian. In the Russian dialect, the entire paradigm has levelled towards **k**, while in Ukrainian, it has levelled towards **ʒ**.

(5) 'I bake'

Standard Russian:	Russian dialect:
p ^h eku	p ^h eku
p ^h effoʃ	p ^h ek ^h oʃ
p ^h effot	p ^h ek ^h ot
p ^h effom	p ^h ek ^h om
p ^h effote	p ^h ek ^h ote
p ^h ekut	p ^h ekut

'I can'

Standard Russian:	Ukrainian:
moʒu	moʒu
moʒeʃ	moʒeʃ
moʒet	moʒe
moʒem	moʒemo
moʒet ^h e	moʒete
mogut	moʒut

Regarding the evolution of a Type-3 pattern from a Type-2 pattern, the 'unconscious and psychical principle' may 'come to the rescue of the . . . alternation by endowing it with a new function. Were it not for this function, the alternation would be destined to irrevocable extinction' (Kruszewski, 1881: 22).

Chaos, as we observe it in the domain of anthropophonic phenomena, is only *temporary*. Everything that was once, but is no longer, absolutely necessary from the anthropophonic view is exposed to the effect of *unconscious, psychical factors . . . which] strive to impose complete order and simplicity on language*. (1881: 20; throughout, italics within quotation marks are in the original)

Baudouin de Courtenay's writings on this same topic were intended as a challenge to the Neogrammarians' proposal that sound change is 'law-governed' in the sense that we can, with sufficient data, predict diachronic endstates:

Between the starting and ending point of historical change (such as the transition from an original **k** to **ŋ**, or **ei** to **i**) there is no relationship that could be interpreted as a law of evolution . . . Any conditioned combination falling under the concept of 'law' belongs to the field of imperceptible [at the time of his writing, D.S.] microscopic differences. Genuine laws of causality are *hidden in the depth*, in the intricate combination of the most diverse elements. (1910: 272, 276)

Baudouin de Courtenay's wrote that the *genuine* law-governed primitives that operate on linguistic patterns derive not from the observation of superficial linguistic patterning, but instead from four main sources: (i) 'the psychological world of the individual' [cognition, D.S.], (ii) 'the biological and physiological world of a given organism' [articulatory phonetics, D.S.], (iii) 'the external, physical world' [acoustic phonetics, D.S.] and (iv) 'the social world (the transmission of linguistically expressed ideas from one individual to another' [psychological matches and mismatches between speaker and hearer, D.S.] (1910: 261): 'The complexity and causes accounting for the emergence and preservation of alternations must ultimately be ascribed to communal life and the physical (anatomico-physiological) and psychological make-up of the members of a speech community.'

2.2 The Post-Kazanians

The interacting pressures of phonetic variation ('the physical'), cognition ('the psychical') and also the inevitable psychological mismatches between speaker

and hearer ('The social world'), is similarly taken up by other nineteenth-century scholars. Consider the writings of Paul (1880: 44), who additionally considers the important role that *frequency and recency of usage* may have on linguistic structure:

... Variability of production, which remains unnoticed because of the narrow limits in which it moves, gives the key to our comprehension of the otherwise incomprehensible fact that a change of usage in the sounds of a language sets in and comes to its fulfillment without the least suspicion on the part of those in whom this change is being carried out.

If the motory sensation were to remain always unchanged as a memory-picture, the insignificant deviations would always centre round the same point with the same maximum of distance. In fact, however, this sensation is the product of all the earlier impressions received in the course of carrying out the movement in question, and, according to a common law, the impressions, not merely those which are absolutely identical, but also those that are imperceptibly different from each other, are fused into one. Correspondingly to their difference, the motory sensation must be somewhat modified with each new impression, to however insignificant an extent. It is, in this process, of importance that the later impressions always have a stronger after-influence than the earlier. It is thus impossible to co-ordinate the sensation with the average of all the impressions during the whole course of life; rather, the numerically-speaking inferior may, by the fact of their freshness, outbalance the weight of the more frequent . . . There thus gradually arises, by adding together all the displacements (which we can hardly imagine small enough) a notable difference . . .

Schuchardt (1885: 57–58) in his challenge to Neogrammarian doctrine writes in similar terms, and specifically implicates *token frequency*, and its interaction with *recent* versus *remote* speech acts, as important factors in certain forms of sound change:

The change of a sound, its progress in a certain direction . . . consists of the sum of microscopic displacements. It is, therefore dependent upon the number of repetitions. If *x* requires 10,000 repetitions to become *x'*, these repetitions are to be counted within individual words, nevertheless. An *x* spoken one time each in 10,000 different words would not become *x'*. I will not deny that a word that has been spoken 10,000 times can favor the development of the sound *x* to *x'* in a word spoken only 8,000 times, etc. The greater or lesser frequency in the use of individual words . . . is . . . of great importance for their phonetic transformation . . . Rarely-used words drag behind; very frequently used ones hurry ahead . . . They have been

compared to small coins that, as they pass from hand to hand rapidly, are soon worn thin.

Such ideas have never died away. Among twentieth-century scholars, consider Hockett's musings on the subject (1958: 443):

If some speaker of English, over a period of years, were to hear a relatively large number of initial /t/'s with unusually inconspicuous aspiration . . . the location of the frequency maximum would drift, and his own speech would undergo the same modification. We would not, of course, expect any *single* speaker of English to have such an experience. In general, individuals who are in constant communication with each other will experience essentially parallel changes in their . . . articulatory habits. It is just this sort of slow drifting about of . . . distributions, shared by people who are in constant communication, that we mean to subsume under the term 'sound change'.

Hockett further elucidates a *wundt-curve*-like model of levelling vis-à-vis frequency of usage (1958: 396–397):

Other things being equal, irregular forms of high frequency are less apt to be replaced than are rarer ones . . . [If] an irregular form is frequently used, a child learning his native language will hear it many times, and may never come out with any analogically produced regular alternant. Even if he does, he probably already knows the inherited irregular form and may reject his own innovation . . . For a rarer irregular form this argument applies in reverse . . . Under some circumstances, extreme rarity may preserve an irregular instead of helping to lose it. The process, however, is quite different. The word *spake* (past tense of *speak*) and *beholden* still occur from time to time; it would seem that the rarity and irregularity of the forms constitute an integral factor in their peculiar archaic flavor, and it is because of the latter that the forms are used.

Martinet (1952) adds an important new ingredient to the general recipe of usage-based phonological change. While readily acknowledging the importance of phonetic and cognitive pressures on patterns of sound change, as well as the effects of frequency of usage, Martinet ascribes special import to the issue of *lexical semantic confusion*: all else being equal, certain diachronic developments – specifically, sound mergers – are more likely to proceed if the *functional load* of the relevant phonological opposition is low. That is, if a given opposition is responsible for a large number of minimal pairs, a merger of the two values is less likely to proceed.

According to Martinet, the tendency towards merger of an opposition is favoured to the extent that (i) The values in opposition are phonetically similar, (ii) The number of minimal morpheme pairs that the opposition is responsible for is low, (iii) The number of minimal pairs within a correlated opposition is low (or the opposition is uncorrelated, where *correlation* refers to the Trubetzkoyan notion of a sound series that is opposed to another by one feature), (iv) The minimal pairs belong to different syntactic categories, (v) The token frequency of one or both members of the minimal pairs is low and (vi) The presence of additional morphological markers serves a disambiguating function.

Martinet is thus moving towards a more holistic functional approach to usage-based phonology, one that, in theory at least, incorporates the role that lexical semantic confusion on the part of the listener might play in the diachronic trajectory of sound systems. As we discuss in Section 2, the role of lexical semantic confusion features prominently in Labov's proposed mechanism of usage-based sound change.

2.3 Boundary Signals and Prosodies

The role of juncture cues should certainly be included in any discussion of usage-based approaches to phonology: aspects of phonological structure can be harnessed by users to assist them in parsing the speech stream into its constituent parts. As Trubetzkoy (1939: 273) notes:

In addition to the phonological means serving to distinguish individual units of meaning (sememes), each language has a number of means that effect the delimitation of such individual units of meaning . . . Each language possesses specific, phonological means that signal the presence or absence of a sentence, word, or morpheme boundary at a specific point in the sound continuum.

Trubetzkoy calls these 'boundary signals', and continues with a helpful analogy:

They can probably be compared to traffic signals . . . It is possible to get along without them: one need only be more careful and more attentive. They, therefore, are found not on every street corner but only on some. Similarly, linguistic delimitative elements generally do not occur in all positions concerned but are found only now and then. The difference lies only in the fact that traffic signals are always present at 'particularly

dangerous' crossings, whereas the distribution of linguistic delimitative elements in most languages seems to be quite accidental. This is probably due to the fact that traffic is artificially and rationally regulated, while language shapes and develops organically.

In all, Trubetzkoy taxonomizes boundary signals by noting that (i) they may be contrast-expressing or contrast-suspending, (ii) they may be positive (cueing a boundary) or negative (cueing a non-boundary), (iii) they may be phonemic or non-phonemic and (iv) they may be individual signals (a single segment) or group signals (a segment sequence).

For example, in Barra Gaelic the aspirated occlusives are found only in word-initial position, and the long vowels, the central vowels and the nasalized vowels are only found in word-initial syllables. These are contrast-expressing boundary signals, rather than contrast-suspending ones, though it must be emphasized that their role as contrast-expressing boundary signals is a consequence of contrast suspension in other positions. Another example: in Japanese **g** occurs only word-initially, and **ŋ** occurs only intervocally (word-medially). Since the two are not responsible for minimal pairs, we are dealing not with a phonemic boundary signal, but rather with a non-phonemic one. Such cases can be multiplied any number of times: elements in complementary distribution, one of which is conditioned by proximity to a boundary, always serve this demarcative function.

Firth's (1948) discussion of prosodies has many parallels to Trubetzkoy's boundary signals. The primary phonological distinction Firth attends to is that between *sounds* and *prosodies*. *Sounds* are components of phonological structure that do not play a syntagmatic role. Sounds occur in *phonematic systems*, and possess solely paradigmatic functional relevance, manifested by 'sound substitutions'. Employing the cover terms **C** and **V**, a phonematic system of sounds may occupy a **C** or **V** position and, as such, sounds function contrastively, but impart no syntagmatic information (apart from their being limited to either a **C** position or a **V** position). *Prosodies*, by contrast, are exactly those elements that *do* impart syntagmatic information. This is not to say that some phonetic value cannot be both a sound *and* a prosody in the same language. In such cases, instances of this value are still regarded as phonologically distinct from each other in contexts where they play distinct – paradigmatic or syntagmatic – roles.

In many ways, prosodies are comparable to Trubetzkoy's boundary signals, though Trubetzkoy adheres to a segmental (or segment-sequential) notion of boundary signals, whereas Firth's prosodies are not comparably limited in shape. Employing conventional terminology for the moment, a prosody may consist of a 'segment'-sized element, a 'sub-segment'-sized element, or a 'suprasegment'-sized element. But it's misleading to relate prosodies (or sounds)

to segments at all, as the prosody-sound distinction is based solely on whether the (sub-) system plays a syntagmatic or paradigmatic role; simply stated, if a value is predictable with respect to its distribution in some domain, it qualifies as a prosody; if a value is not predictable with respect to its distribution in some domain, it is a sound. And though Firth does indeed talk in terms of consonants and vowels, it is clear from his exposition that these are mere terminological expedients.

Robins (1957: 192) attempts to elucidate the sound-prosody distinction:

Phonematic units refer to those features or aspects of the phonic material which are best regarded as referable to minimal segments, having serial order in relation to each other in other structures. In the most general terms such units constitute the consonant and vowel elements or C and V units of a phonological structure. Structures are not, however, completely stated in these terms; a great part, sometimes the greater part, of the phonic material is referable to prosodies, which are, by definition, of more than one segment in scope or domain of relevance, and may in fact belong to structures of any length . . . A structure will thus be stated as a syntagmatic entity comprising phonematic or segmental units and one or more prosodies belonging to the structure as a whole.

Robins (1957: 192–193) emphasizes that the phonetic exponents of prosodies need not pervade their domain of association: as already noted, a prosody may be ‘segmental’, ‘sub-segmental’, or ‘suprasegmental’ in its phonetic exponent, its status as a prosody being a consequence of its predictable distribution within some domain:

Broadly speaking [prosodies] come about in two ways. (1) In the first case a feature may be spread or realized phonetically over a structure, such as a syllable, as a whole . . . (2) In the second case may be mentioned features which are not realized phonetically over the whole or large part of a structure, but which nevertheless serve to delimit it, wholly or partly, from preceding or following structures, thus entering into syntagmatic relations with what goes before or after in the stream of speech. By virtue of their syntagmatic relations in structures, such features may be treated as prosodies of the structures they help to mark or delimit . . . (indices added)

As discussed in Section 2, modern investigations of the functional value of boundary signals and certain aspects of prosodies fall under the rubric of so-called *transitional probabilities*.

3. Modern Currents in Usage-Based Phonology

The flame carried by nineteenth-century scholars who may broadly be considered 'usage-based phonologists' dimmed to a mere flicker in the post-war period. Nonetheless, usage-based approaches have survived and, starting with the pioneering scholarship of Labov and Ohala in the 1970s and 1980s, have begun to flourish once again. Implicit in both these research programmes is a role for exemplar modelling, discussed in detail by Bybee, among others. We turn to this issue first.

3.1 Exemplar Modelling

Bybee's usage-based approach to phonological structure (e.g. 2001, 2006a and b) has been greatly influenced by her nineteenth-century predecessors, especially, Kruszewski and Schuchardt. According to Bybee, linguistic categories and their clumping into larger units emerge as a consequence of patterns' frequency of occurrence and co-occurrence. When elements frequently pattern together, they are likely to emerge as independent functional units of language. Many sound changes are the result of phonetic processes that apply as a consequence of actual language use, and consequently, those words that are used more frequently are more likely to undergo phonetic processes. This is exactly the proposal of Schuchardt, over 100 years previous.

Bybee (2001) provides many case studies – most from English, Spanish as well as a detailed discussion of French liaison – illustrating how sound changes may begin with words and phrases of the highest frequency, and then may gradually diffuse through the lexicon. For example, whereas frequent words like *camera* and *every* have lost their medial schwas, less common words with comparable structure like *mammary* and *homily* retain these schwas. While frequent words are more likely to lead the way in certain phonetic reductions and assimilations, they are also more likely to resist levelling processes. For example, high frequency strong verbs like *kept* have resisted the regularization that has affected less frequent past tense forms such as *wept* → *weeped*, exactly the scenario presented by Hockett.

Bybee proposes that the lexicon is fully specified with phonetic detail, and is highly structured with interconnections among phonetically and semantically parallel structures. The more similar that lexical entries are in terms of their structural properties, then (i) the more likely that the morphological structures of these words will emerge and (ii) the more likely that the words will be subject to the same phonological processes.

Speaker knowledge of phonotactic regularities is claimed to be an emergent consequence of frequency of type occurrence. Bybee cites studies which indeed

show that listeners' acceptability of sound sequences that are embedded in nonce forms correlates highly with these sequences' type-frequency in real words, and with their overall similarity to real words. Acceptability judgements here are gradient, showing that more familiar structures are more acceptable to listeners, and less familiar structures are less acceptable. Comparable work on Arabic by Frisch and Zawaydeh (2001), Frisch (2004), and Frisch et al. (2004) is fully consistent with Bybee's findings: speakers possess knowledge of the phonotactic regularities of the language, knowledge that is statistically nuanced in the sense that speakers can make gradient judgements on the 'naturalness' of nonce forms that parallel the prevalence of such patterns in their lexicons.

Such proposals support a specifically *exemplar*, *episodic*, or *multiple-trace* approach to lexical organization. First introduced to phonology by Johnson (1997), exemplar modelling has its origins in the classic categorization study of Shepard et al. (1961), which in turn influenced a number of further important studies on categorization of similar and dissimilar sensory items, among them Tversky (1977) Tversky and Gati (1978, 1982), Medin and Schaffer (1978), Medin (1983) and Gluck and Bower (1988). These researchers observe that items may be regarded as more similar or less similar to each other based not only on their physical attributes, but also on the contexts in which items are placed, and the functional role that items play.

The basic proposal of exemplar theory is that categorization proceeds from *experience* with actual sensory objects: perceptual categories emerge from *repeated exposure* to similar sensory events, where, as just noted, similarity is not determined solely on physical grounds, but also by the context in which items are placed, and the functions to which items are put. In general, the more often a sensory event is perceived, the more likely it will come to emerge as a categorical component of the system. Nosofsky (1986, 1988) and Goldinger (1997, 1998) further explore the role of categorization within a specifically exemplar model of categorization. Nosofsky proposes that perceptual stimuli are categorized based on their degree of similarity to stored exemplars. Goldinger suggests that an 'episodic' model of word learning (and memory in general) obviates the need for learners to match perceived speech to idealized templates or prototypes in a normalization procedure. He discusses a number of lines of evidence supporting the claim that humans have a remarkable memory capacity, one that is capable of storing richly detailed information about both linguistic and nonlinguistic perceptual stimuli.

The application of these ideas to phonology seems an obvious next step: surely, allophonic relatedness presents a scenario in which physical distinctness is overridden by the functional role that the objects play in the linguistic system (Silverman, 2006a). In phonology, the sensory events in question are speech tokens, and the categories that may emerge are those components of the speech stream that are repeated over and over again. In this approach,

phonological categories are the emergent consequence of language *use*. Bybee (2006a: 717) discusses several advantages of the exemplar approach:

- (1) Exemplar representations allow specific information about instances of use to be retained in representation,
- (2) Exemplar representations provide a natural way to allow frequency of use to determine the strength of exemplars and
- (3) Exemplar clusters are categories that exhibit prototype effects. They are organized in terms of members that are more or less central to the category, rather than in terms of categorical features.

Note that Bybee's approach need not embrace the segment as a phonological primitive. Given that repeated patterns are of many shapes and sizes, the phonological units that might emerge may consist of articulatory routines of varying length and complexity. In Japanese for example, the single tongue blade gesture in the sequence *ʃi* is argued to historically derive from *si*, which is claimed to have involved a sequence of blade gestures. Due to the frequency of their co-occurrence, this gestural sequence gradually merged in terms of tongue position, culminating in the single articulatory gesture in evidence today.

Since the pioneering work of Johnson, the exemplar model has been harnessed to varying degrees of rigor by a number of researchers, including Steels (2000), Silverman (2000, 2006a and b), de Boer (2001), Pierrehumbert (2001), Liberman (2002), Wedel (2004, 2006), Yu (2004, 2007), Plug (2005, 2010), Ernestus and Baayen (2006). Steels and de Boer, for example, computationally model vowel systems as self-organized complex dynamic systems; Plug investigates certain discourse patterns of phonetic reduction from an exemplar theoretic and usage-based perspective. The exemplar model has perhaps been most compellingly applied by Labov (1994).

3.2 Semantic Misperception

Labov's proposed mechanism of sound change (discussed in detail in Labov, 1994) is firmly exemplar theoretic in orientation, and furthermore, is a direct descendent of Martinet's functional account, though applied with a great deal more rigor. Consider first an example case of a shift in usage: in French, plural *s* has been lost (except when a vowel follows), and thus, for example, the plural article (earlier, **las* in all contexts) runs the risk of being homophonous with the singular, that is, *la*. However, the plural is now (usually) signalled by a change in vowel quality: **las* → *le*. As Labov asserts, '[This] show[s] how long-range changes in the French phonological, morphological, and syntactic systems compensated for sound changes, in ways that suggest a causal link' (1994: 570).

Comparable patterns exist in any number of systems, including Boston Puerto Rican Spanish (Hochberg, 1986): plural **s** is variably absent, but its absence is more often encountered in inherent plurals, and less often encountered when the loss of **s** would result in semantic ambiguity, thus **las plantas** (*the plants*; cf. **la planta** *the plant*) but **muŋa planta** (*many plants*). Consider further the pattern in (6). Note in particular the distinction between 2nd sg and 3rd sg verb agreement; 2nd person is marked by **s**, whereas 3rd sg lacks this **s**.

(6) 'to study'

	Sg	Pl
1st	(jo) estudio	(nosotros) estudiamos
2nd	(tu) <u>estudias</u>	
3rd	(él, ella, usted) estudia	(ellos, ellas, ustedes) estudian

The **s** in the 2nd sg (underlined) should delete the least often, since it is the sole marker of the 2nd–3rd contrast. In fact, the **s** in the 2nd sg drops more often than the overall average. However, pronoun use increases in the context of this **s**-drop, thus morphologically salvaging the phonologically neutralized contrast. Interestingly, among educated Madrid speakers, in which **s** is not undergoing attrition, pronoun use is significantly lower. Pronoun use increases upon switch-reference, however (when a new subject is introduced) (Cameron, 1992).

As Labov writes, 'If speakers do not consciously or unconsciously adjust their sentences to maximize the transmission of meaning, then we need to find some other mechanism that accounts for the systemic adjustments that maintain informational content' (1994: 585). His proposed mechanism is *probability matching*: animals (including humans) have the capacity to replicate observed frequencies of events in their behavioural responses. Studies indicate that certain of these behaviours must be the result of perception and calculation rather than reward, since animals may adjust their behaviour even without having been actually reinforced. So-called variable rule learning may thus easily be seen as proceeding in the same fashion: the statistical distribution of speech tokens within the phonetic space is calculated by language learners, and, most remarkably, is largely matched in their own speech productions.

Labov applies probability matching and exemplar modelling to aspects of sound change. As stated, the basic idea is that language users are especially adept at matching in their own productions the variation that they perceive, such that variation is conventionalized in the speech community. However, as Labov writes (1994: 586):

It is not the desire to be understood, but rather the consequence of misunderstanding that influences language change. This mechanism

implies a mismatch between producer and interpreter: the type of built-in instability that we would expect to find behind long-term shifts in language behaviour.

For example, if the word *drop* is produced as **dræp** it might be understood correctly, since it is unlikely to be confused with another word (there is no English word *drap*). No matter how small such an effect, repetition may come to shift the pool of exemplars in terms of its phonetic properties. Consequently, such interlocutions may, over time, lead to an overall fronting of the low back vowel. Alternatively, **dræp** might not be understood, and the token may simply be thrown out, having no effect on the exemplar pool, thus inhibiting any change. In a fashion comparable to this latter scenario, if a token of *block* is produced as **blæk**, there is a greater likelihood of misunderstanding since both *block* and *black* are actual words. Again, the role of these misunderstandings is to *inhibit* sound change, since such tokens will not be pooled with the listener's store of exemplars for the word *block*. The contrastive vowel qualities – as a passive consequence of language use – may enjoy a comfortable perceptual buffer zone.

3.3 Phonetic Misperception

Like several nineteenth- and early twentieth-century scholars before him, Ohala (e.g. 1981, 1983, 1989, 1993) suggests that variation in speech is the fodder for many sound changes. Harnessing modern experimental techniques, Ohala expands upon these earlier proposals in his use of the laboratory to show how certain phonetic (pre-) conditions may give rise to particular sound changes. Like Labov, he gets much mileage out of the proposal that listeners, as opposed to speakers, are progenitors of sound change, thus harkening back to Baudouin de Courtenay's (1910) discussion of *lapsis auris*. Unlike Labov, however, for Ohala, the locus of listener misperception is phonetic, rather than semantic.

According to Ohala, four major scenarios may play themselves out over time as a consequence of the interplay between the acoustic signals that speakers produce and the interpretations of these signals by listeners: (i) correction of acoustically unclear signals, resulting in diachronic stability, that is, no sound change, (ii) confusion of acoustically similar sounds, (iii) hypo-correction and (iv) hyper-correction.

Hypo-correction involves listeners interpreting a context-dependent phonetic effect (often coarticulatory or assimilatory) as context independent. For example, the nasalization present on vowels in the context of a following nasal consonant may be ('mis-')attributed to the vowel, rather than as a context-dependent

feature of the nasal consonant. Listeners may thus fail to correct for a predictable feature of the speech signal. Once a listener interprets the nasalization on the vowel as primary, the nasal consonant itself may be interpreted as context dependent, in time withering to zero, thus, $VN > \check{V}N > \check{V}$. Note that the end-point of such a sound change is incipient in its starting point, in the sense that low-level phonetic variation involving vowel nasalization provides the necessary fodder for the initiation of the change.

Regarding hyper-correction, listeners may over-correct a component of the speech signal, misinterpreting a context-independent property as context dependent. For example, in Latin **kwin̄kwe:** > **kin̄kwe:**, assuming a degree of labiality persists through the first vowel, a listener may mistakenly conclude that the labiality of the first vowel is simply an automatic ‘spillover’ from the second velar release, so they ‘undo’ it, (‘mis-’)attributing it solely to the second **k**. The result of such a hyper-corrective sound change is dissimilation.

Though not overtly discussed by Ohala, this notion of ‘misattribution’ on the part of listeners is predicated on a specifically segmental approach to phonological structure. Blevins’ approach (2004, 2006a and b), essentially the same as Ohala’s, overtly embraces the segment as a phonological primitive, typologizing certain sound changes as the result of ‘ambiguous segmentation’ due to (i) change, (ii) chance or (iii) choice.

Sound change due to *change* involves the phonetic signal being ‘misheard by the listener due to perceptual similarities of the actual utterance with the perceived utterance’ (2004: 32). For example, **anpa** may be misheard as **ampa**.

Sound change due to *chance* involves a phonetic signal that is ‘accurately perceived by the listener but is intrinsically phonologically ambiguous, and [so] the listener associates a phonological form with the utterance which differs from the phonological form in the speaker’s grammar’ (2004: 32). For example, a speaker may say **?a?**, and a listener may recover the signal accurately, but does not faithfully reproduce the speaker’s mental representation of the utterance: the listener constructs **/a?/** where the speaker constructs **/?a/**. Clearly, like Ohala before her, Blevins assumes the existence of segments, and also the existence of the generative-theoretic notion of underlying representations (Chomsky and Halle, 1968).

Sound change due to *choice* is characterized thus:

Multiple phonetic signals representing variants of a single phonological form are accurately perceived by the listener, and due to this variation, the listener (a) acquires a prototype or best exemplar of a phonetic category which differs from that of the speaker and/or (b) associates a phonological form with the set of variants which differs from the phonological form in the speaker’s grammar. (Blevins, 2004: 33)

For example, a speaker may say **ka'kata kǎ'kata kkata** for /**kakata**/, while the listener hears **ka'kata kǎ'kata kkata** but mentally constructs /**kkata**/. Thus, *choice* too crucially relies on the notion of underlying representations.

3.4 The Listener's Role: Interpreting Speaker Intent, or Matching Speaker Behaviour?

Ohala and Labov clearly entertain different hypotheses regarding the listener's role in certain aspects of sound change. Recall that, in general terms, Ohala proposes that listeners are intent on interpreting the *phonetic* intentions of speakers, and that certain types of sound change are a consequence of listeners' sporadic 'incorrect' conclusions about these phonetic intentions. Labov, by contrast, proposes that listeners are, rather, exceptionally talented in interpreting the phonetic signal produced by speakers, as evidenced by the fact that they are able to match the very variation they perceive, in a form of probability matching. For Labov then, certain sorts of sound change may be a consequence not of listeners' sporadic misinterpretation of the *phonetic* signal, but rather, a consequence of listeners' sporadic misinterpretation of the *semantic* content that rides on this phonetic signal.

Consider the findings of Öhman (1966) and Manuel (1990, 1999) in light of these two competing accounts. These authors investigate patterns of coarticulation: Öhman investigates cross-linguistic patterns of vowel-to-vowel coarticulation (**VCV**) through intervening consonants, and Manuel investigates cross-linguistic patterns of vowel coarticulation due to consonantal context (**CVC**). Both find that different languages possess different patterns of coarticulation in these contexts, and further, that at least a certain amount of the observed language-to-language difference in coarticulation may be attributable to the language-particular system of contrastive values.

Öhman considers **VCV** coarticulation in Swedish, English and Russian. This last language, unlike the first two, has a series of palatalized consonants that may influence an observed curtailment of the degree of coarticulation, such that palatal contrasts are recoverable in the speech signal:

In Swedish and English, the stop consonants seem to coarticulate relatively freely with the vowels . . . there are languages, such as Russian, in which the instructions for the stop consonants are made . . . as in English or Swedish but with the additional feature that the vowel channel must simultaneously receive exactly one of two fixed commands [palatalization or velarization]. (1966: 166)

The data discussed by Manuel (1990) and Manuel and Krakow (1984) are also consistent with the idea that coarticulation is influenced at least in part by the distribution of contrastive values in the phonetic space, and that coarticulation may be curtailed to the extent that it (at least sometimes) does not jeopardize these contrastive values. Manuel and Krakow find that there are larger V-to-V coarticulation effects in languages with smaller vowel systems, and smaller coarticulatory effects in languages with larger vowel systems. For example, Shona and Swahili, with 5-vowel systems, may display more vowel coarticulation than a language like English: on the one hand, because the vowels in English are more crowded in the articulatory/acoustic space, the range of production for each one would be rather small so as to maintain distinctions among them; on the other hand, as the vowel qualities of Shona and Swahili are fewer, they could presumably tolerate larger ranges of production without running the risk of encroaching on each other's distinctive spaces.

These studies, along with quite a few others that investigate system-influenced patterns of conventionalized (co-) articulatory routines (among them Clumeck, 1976; Beddor et al., 1986; Recasens, 1987; Recasens et al., 1998; Beddor and Krakow, 1999; Beddor et al., 2002) are consistent with Labov's proposal that variation in speech is conventionalized within speech communities: the fact that coarticulation is limited in just those contexts where lexical contrasts would otherwise be jeopardized is readily explainable in the diachronic scenarios envisioned within a Labovian account. Such patterns lend themselves less readily to an account in which listeners are formulating hypotheses about the phonetic intentions of speakers. Under this latter account, it is either a pure coincidence, or is rather due to a circuitous chain of events, that speech variation is conventionalized on a language-to-language basis in ways that bear the clear mark of lexical semantic pressure.

Moreover, assuming an Ohalaian, phonetically-based approach to speaker-listener mismatches as a factor in sound and language change, it would be purely coincidental that aspects of morphological variation pattern in comparable ways to aspects of variable phonetic patterning. Indeed, probability matching in language is found in domains that are surely not explicable in the phonetic terms proposed by Ohala, including variable morpheme usage both in real-world settings (see, for example, Poplack, 1980a and b), and also in laboratory settings. For example, Hudson and Newport (1999) performed an experiment in which subjects were exposed to variable patterns in a contrived mini-language during a learning phase, and came to reproduce this variation in their own speech patterns during a testing phase. In this study, nouns were variably marked with a determiner. Subjects were divided into groups, which differed in the extent to which the nouns they heard possessed this marker: one group was exposed to nouns, 75% of which had the marker, and another group was exposed to nouns, 25% of which had the marker. In the testing phase,

subjects largely matched their usage to their exposure. That is, subjects in the 75% group produced about 75% of their nouns with the marker, and subjects in the 25% group produced about 25% of their nouns with the marker.

All such findings are consistent with the proposal that certain so-called low-level or phonetic effects may in fact be the result of deep, systemic pressures many times removed from the physical systems that proximally underlie speech.

Note finally that a corollary to Labov's 'semantic misperception' approach offers a compelling account of the observed link between token frequency and articulatory simplifications: it is exactly because certain words are frequently encountered in the speech stream that they are more predictably present. Because of their constant repetition and their consequent predictability, those particular spontaneous variants that are slightly simplified may yet effectively convey the intended meaning to listeners. Due to probability matching, in time, these simplifications may become conventionalized. Thus, accurate semantic perception proceeds despite phonetic simplification (Silverman, 2006a, 2010, in prep.).

3.5 Near Merger and Near Neutralization

Labov's discovery of *near mergers* and *near neutralizations* (e.g. Labov, 1966; Labov et al., 1972; Labov et al. 1991) has had great success in explaining many previously ill-understood sound changes, and has inspired a significant amount of work as well. (See, for example, Dinnsen and Charles-Luce, 1984, and Charles-Luce, 1993 on Catalan; Port and O'Dell, 1985, and Port and Crawford, 1989 on German; Slowiaczek and Dinnsen, 1985 on Polish; Pye, 1986 on Russian; Warner et al., 2004 on Dutch; Gerfen and Hall, 2001, and Bishop, 2007 on Andalusian Spanish.) Near merger occurs when there is significant token-to-token phonetic overlap of two (or more) phonological values, such that language users may not be aware of the phonetic distinction that is variably in place. Note that listeners are clearly sensitive to these values' nearly merged status, since they recapitulate the pattern in their own speech (in a form of probability matching), but they may lack conscious awareness of their persistent small degree of difference. (Indeed, it emerges as a corollary to a specifically usage-based phonology that *speaker intuitions* should perhaps play no role whatsoever in linguists' proposals about the structural components of language.)

If we assume that genuinely merged values cannot be undone – unmerged – by linguistic means (this is Garde's Principle, after Paul Garde, 1961), then the existence of near mergers offers a compelling explanation for patterns that have been (mistakenly) analysed as having merged in the past, only to unmerge at

a later point in time. The doctrine of Uniformitarianism (originally applied to geological strata) states that the laws governing the patterning of natural phenomena are equally valid across all space and time, and thus, 'knowledge of processes that operated in the past can be inferred by observing ongoing processes in the present' (Christy, 1983). Consequently, as Labov (1974) writes, we might use the present to explain the past. The existence of near mergers today is good evidence for their existence in the past. More particularly, if we find near mergers in the present in exactly those cases that purportedly underwent complete merger in the past, then we may conclude that the values did not, in fact, completely merge at all in the past. Rather, they merely nearly merged, and this near merger has persisted to the present.

The contemporary evidence for near mergers may thus provide a compelling explanation for putative cases of historic 'unmergings'. For example, in Middle English *meet*, *meat*, and *mate* possessed distinct vowel qualities: *e*, *æ*, *a*. During the sixteenth century, the vowels *æ* and *a* purportedly merged towards *ɛ*, but during the seventeenth century they purportedly unmerged, with (historic) *æ* and *e* merging towards *i* (and historic *a* rising to *e*). Labov reports that, in fact, certain contemporary Belfast dialects possess the near merger of *æ* and *a* (Milroy and Harris, 1980). That is, the vowel qualities that purportedly merged – and then purportedly unmerged – in the past are, in fact, nearly merged today. Labov proposes that the values never really merged at all. Instead, they engaged in a near merger that, quite remarkably, has persisted for several hundred years:

The overlap [in the distribution of the two vowel qualities] has not prevented the distinction between the two classes from being maintained for almost three hundred years . . . It follows that speakers are capable of tracing the frequency of occurrence of the two classes . . . and that this differential distribution is a part of their fundamental knowledge of the language . . .

This is a compelling instance of long-term probability matching in language use.

Some of Labov's most famous research investigates near mergers in North American English. For example, consider New York *source* and *sauce*. In so-called r-less dialects, the non-prevocalic **r** found in other dialects typically corresponds to a schwa-like offglide here. Since these same dialects possess **ɔə** in words like *sauce*, the pronunciation of these two words – *source* and *sauce* – is nearly identical, their meager difference more often encountered in recitation speech, less often in spontaneous speech (Labov et al., 1972).

Another case: in Albuquerque, a high school student nearly merged the vowels in *fool* and *full*. Despite a slight though persistent difference in their

phonetic properties, this student felt that all the relevant words possessed but a single vowel quality. He was recorded reciting a list of *fool-full*, *pool-pull* words. When this recording was played to speakers who possessed a better separation of the vowel qualities, they correctly identified the words 83% of the time (Labov et al., 1972).

Charles-Luce (1993) reports on a study of a related phenomenon, near neutralization in Catalan. Her results show that the tendency towards neutralization is indeed affected by semantic factors, just as suggested by Labov's 'consequences of misunderstanding' proposal: 'The perception and production of spoken words is affected differentially by the presence and absence of higher levels of linguistic information and . . . the degree of precision of articulation is inversely proportional to the presence of semantic information' (1993: 29). She finds that a Catalan voicing alternation is more likely to be nearly neutralized (as opposed to completely neutralized) in contexts that would otherwise be semantically ambiguous. As Charles-Luce concludes, 'There may be some on-line assessment by the speaker as to the degree of biasing information present [that] may be quite automatic and learned through experience . . .' (1993: 41).

Charles-Luce's findings may also be viewed as supporting Martinet's earlier claims regarding pressures that might militate against merger: as an emergent outcome of sporadic semantic misinterpretation, there may, under certain conditions, be a passive social pressure against values' merging and neutralizing.

Many excellent longitudinal and latitudinal studies have been inspired by Labov's pioneering work in quantitative sociolinguistics, among them Poplack (1980a and b) on Caribbean Spanish, Eckert (1988) and Guy (1991). Docherty, Foulkes and colleagues (e.g. Docherty and Foulkes, 1999; Docherty et al., 2006; Foulkes and Docherty, 2007; Foulkes et al., 2010) have presented detailed sociophonetic investigations of Tyneside English among other dialects, focusing in particular on so-called socially structured variation. These authors consider many aspects of speech variation that are partially delineated by social setting, including social class differences, age-based differences, sex differences, child-directed speech etc.

3.6 Natural Selection as Metaphor

Silverman (2006a) is directly inspired by Labov's work on probability matching, Ohala's proposals regarding the phonetic preconditions for sound change, and Martinet's proposals regarding the role of functional load in the tendency towards merger. In keeping with his strictly functional approach to phonology, Silverman typologizes synchronic *sound substitutions* into three logical/functional categories: (i) *contrastive* (meaning-changing), (ii) *neutralizing* (which are

re-defined as exclusively homophone-inducing alternations) and (iii) *allophonic* (meaning-preserving). Employing an exemplar-theoretic approach to lexical organization, Silverman argues that morphemes are not broken down by language users into smaller-sized sound units unless there is evidence from alternation to do so. He applies Darwin-like evolutionary principles to patterns of sound change, proposing – like Kruszewski and Baudouin de Courtenay before him – that the explanation for synchronic patterns of usage resides in phonetic and functional pressures that interact across generations of language use.

The basic components of his approach are Darwin-inspired, including (i) speech variation (cf. mutations), (ii) communication from speaker to listener (cf. reproduction) and (iii) the increased likelihood of semantically unambiguous speech tokens being stored and recycled as listeners become speakers (cf. natural selection). Recall that, as discussed by Labov, variation in speech can sometimes lead to semantic confusion for listeners (as when one word is confusable with another due to their phonetic similarity), and that this confusion may, over generations of speakers, lead to the better separation of phonological categories. Under this view (as in the Kazan School) allophonic alternants are viewed as the culmination of a series of small, natural changes to the system that takes place over generations of speakers. In more recent work (Silverman, 2010, in prep.), he investigates the proposal that neutralizing patterns (according to the term's traditional definition) are tolerated to the extent that they do not derive excessive homophony, à la Martinet and Charles-Luce.

Wedel (2006) also assumes this strong version of functional, Darwin-styled phonology, discussing three pressures on sound change that derive directly from theories of evolutionary biology: (i) *pruning of lines of inheritance*, which involves the slow memory decay of individual speech tokens – including outliers – and their subsequent replacement by more recent tokens, (ii) *blending inheritance*, which involves the averaging of multi-modal distributions, resulting in a winnowed uni-modal distribution and (iii) *natural selection* (much like Silverman's proposal). Wedel runs various computer simulations that demonstrate how each of these pressures may lead to a naturalistic distribution of sound categories in the perceptual space.

3.7 Transitional Probabilities

Trubetzkoy's and Firth's work on the functional relevance of boundary signals and prosodies, respectively – that is, their role in serving as an aid to parsing – has, in recent years, been experimentally studied by a number of scholars who are focusing on the functional value of so-called *transitional probabilities*

(e.g. Saffran et al., 1996a and b; Aslin et al., 1998). These scholars investigate the utility of transitional probabilities in both adult and infant learning of contrived mini-languages, finding that, indeed, statistically rare sound sequences found at 'word' boundaries (of course, in these experiments they are not real words) serve to cue these boundaries. The necessary flipside to this finding is that statistically more prevalent sound sequences – those involving neutralization or contrast suspension within some domain – may function as negative boundary signals, that is, they may cue a non-boundary. Saffran et al. (1996b: 609) provide a nice cross-modality illustration of what they intend to investigate:

One might discover words in the linguistic input in much the same way that one discovers objects in the visual environment via motion: the spatial-temporal correlations between the different parts of the moving object will be stronger than those between the moving object and the surrounding visual environment.

Formulaically, the transitional probability of **y** given **x** is shown in (7):

$$(7) \frac{\text{frequency of pair } \mathbf{xy}}{\text{frequency of } \mathbf{x}}$$

If this ratio is high, the presence of **x** is a good predictor of a following **y**; such sequences might thus serve as negative boundary signals – Trubetzkoy's 'green light' – increasing the likelihood that the sequence is word-internal. However, if this ratio is low, then the sequence **xy** may serve as a positive boundary signal.

In one of their experiments, adult subjects were taught a contrived mini-language consisting of four consonants (**p b t d**) and three vowels (**a i u**). Twelve CV syllables were constructed, which were strung into tri-syllabic sequences constituting the 'words' of the language, for example, **bapubi**, **dutabi** etc. Transitional probabilities at 'word' boundaries were lower than transitional probabilities within 'words'. After a training period, subjects were able to determine word structure at levels significantly better than chance, thus showing that they perform complex statistical calculations over the sound sequences they were trained on, even absent semantic feedback.

4. Conclusion

With regard to the matter of transitional probabilities and language development (be it 'ontogenetic' or 'phylogenetic'), it should be emphasized that there is no chicken-or-egg problem: there is no reason – as a matter of principle – to weigh in on the issue of whether, say, language happens to possess certain

incidentally beneficial phonotactic patterns, and learners come to take note of them as they begin to parse the speech stream (i.e. phonotactic patterning drives parsing), or whether the learning process actually comes to shape the phonotactic regularities in a way that makes it easier to parse (i.e. learning drives phonotactic patterning). Rather, the complex array of linguistic subsystems is subject to specifically *co*-evolutionary pressures: the manifold evolutionary pressures on the linguistic system – among them, the statistical analytic proclivities of learners, and the natural phonetic pressures that may come to limit the phonetic shapes of words – cannot be cleaved. A usage-based phonology thus entails no proscriptions against the so-called mixing of (linguistic) levels that often accompanies non-usage-based approaches (among them, American structuralist and generative approaches). Properly treating language as a ‘complex adaptive system’ (Steels, 2000) that passively evolves as function of its use, the language analyst should not – as a matter of principle – extract one component pressure on language structure to the exclusion of others with which it is necessarily intertwined. Indeed, a usage-based phonology overtly embraces the proposals – again, as a matter of principle – that (i) language structure is shaped and changed by conventions of language use and disuse, and that (ii) conventions of language use and disuse are shaped and changed by language structure.

Darwin’s famous ‘tangled bank’ passage would seem appropriate to ponder at this juncture:

It is interesting to contemplate a tangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent upon each other in so complex a manner, have all been produced by laws acting around us. These laws, taken in the largest sense, being Growth with reproduction; Inheritance which is almost implied by reproduction; Variability from the indirect and direct action of the conditions of life, and from use and disuse; a Ratio of Increase so high as to lead to a Struggle for Life, and as a consequence to Natural Selection, entailing Divergence of Character and the Extinction of less improved forms. Thus, from the war of nature, from famine and death, the most exalted object which we are capable of conceiving, namely, the production of the higher animals, directly follows. There is grandeur in this view of life, with its several powers, having been originally breathed by the Creator into a few forms or into one; and that, whilst this planet has gone circling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being evolved.

Indeed, it should not be surprising that Darwin's theory of evolution by natural selection has served as a catalyst – either directly or indirectly – for so much research in usage-based phonology, both in the immediate post-*Origin* era, and continuing up to the present day, when technological advances allow for the computational modelling of the self-organizing aspects of complex dynamic systems like language (e.g. Liljencrants and Lindblom, 1972; Lindblom et al., 1984; Steels, 2000; de Boer, 2001; Liberman, 2002; Wedel, 2004, 2006).

There is little doubt that research in speech and language analysis, due to ever-improving technology, will continue to branch and diversify in new and innovative ways that are directly inspired by the Darwin revolution. This author, for one, is quite confident that such research will provide increasingly compelling theories about – and increasingly compelling evidence for – the divergent though intertwined aspects of phonological structure introduced in the opening to this chapter, that is, that phonological systems consist of discrete psychological categories, and that phonological categories emerge from variable speech tokens.

Part IV

HISTORY OF THE FIELD

This page intentionally left blank

16 Issues in the Development of Generative Phonology

Tobias Scheer

Chapter Overview

1. Introduction	397
2. Normal Science: SPE	398
3. Revolution and Revision	404
4. (Unintended) Counterrevolution: Autosegmental Representations	412
5. Second Revolution: Anti-Derivationalism	422
6. Counterrevolution #2: Neo-Serialism and Representations	439
7. Conclusion	442
8. Notes	444

1. Introduction

It is an ambitious undertaking to write about the history of a field in a chapter length article. This chapter thus only provides an overview that concentrates on a small number of central issues and salient steps in the evolution of generative phonology. With respect to the time scale covered in the chapter, the decision to begin with SPE (The Sound Pattern of English) reflects space limitations rather than a judgement on the important developments in the 1950s and 1960s. There is however a growing body of specialized and well-documented literature on the history of (generative) linguistics in general and on phonology in particular that inquires on individual pieces of the overall picture and sometimes covers larger periods.¹

The following pages thus offer a selection based on personal choice, rather than historiography. Nonetheless, the ambition of the chapter follows the traditional motivation of historians: to learn from the past in order to understand the present, and to shape the future. The kind of history that is told is thus not neutral or impartial: it is goal-oriented and functional in the sense that it aims

at isolating important issues, key questions, central ideas, circular movements and real scientific progress.

In this respect, the following topics are discussed: overgeneration, the abstractness debate, (autosegmental) representations and their balance with computation, the nature of computation (rule-based versus constraint-based, serialism versus parallelism) and restrictions on computation (cyclicity). These issues and the theories from which they emerged will be located with respect to the different evolutionary strands, and with respect to each other.

Finally, the chapter is based on published (or unpublished, but written) scholarly work while the sociology of the field, which may well contribute to our understanding of the overall picture, is not considered.

2. Normal Science: SPE

2.1 The Generative Architecture: Modularity and Translation

As indicated by the section title, the pages below follow Kuhn's (1962) analytic tool for the evolution of scientific theories:² first generation generative phonology was established during the 1960s and incarnated as SPE in 1968. It was then practiced for a decade or so with the status of what Kuhn calls normal science, that is, following the canonical rules established without calling them into question even though irritating or frankly conflicting evidence was available and well known. In the case of SPE, its Achilles' heel was made explicit by Chomsky and Halle themselves in the famous ninth chapter of SPE, to which we will come back below.

Since normal science was practiced and largely spread before the actual publication of SPE (individual chapters were circulated in the 1960s), both the revolutionary and the revisionist movement that it aroused could be launched in 1968, or shortly thereafter. The revolutionary attack that set out to overthrow the system as such was led by the Natural Phonologies (which originate in David Stampe's 1972 Ph.D), while the revisionist perspective was opened by Paul Kiparsky's (1968 [1973]) manuscript *How Abstract is Phonology?*

The communist and socialist enterprises are the subject matter of Section 3. Before we turn to them, let us look at some key properties of SPE, and namely at those which have set the generative standards and are still in place today against all odds and subsequent evolutions.

SPE makes explicit the architecture of generative grammar that was heralded in *Aspects* (Chomsky, 1965). This architecture is a direct expression of the frame that the then emerging Cognitive Science defines for cognitive activity (the cognitive revolution of the 1950s, see Gardner, 1985 for an overview): language as such is a module (in the sense of Fodor, 1983, that is, a specialized and autonomous

unit in the cognitive system that carries out computation), but its internal structure is modular as well.³ The following quote documents that generative phonology is conceived of as a cognitive module, that is, a computational unit that modifies an input according to a pre-defined set of instructions that applies ‘mechanically’, that is, without taking into account any external factors (computation is autistic or, in CogSci terminology, encapsulated):

The rules of the grammar operate in a mechanical fashion; one may think of them as instructions that might be given to a mindless robot, incapable of exercising any judgment or imagination in their application. Any ambiguity or inexplicitness in the statement of rules must in principle be eliminated, since the receiver of the instructions is assumed to be incapable of using intelligence to fill in gaps or to correct errors. (Chomsky and Halle, 1968: 60)

The modular architecture of the generative grammar is embodied in the so-called inverted T model that is depicted in Figure 16.1. While the inverted T was first introduced in *Aspects* (Chomsky, 1965: 15ff.), the modalities of communication between morpho-syntax and phonology were defined in a book on phonology, *SPE*, rather than in work on syntax. This phonological bias also runs through further evolution: theories of the interface between morpho-syntax and phonology are phonological theories made by phonologists on the grounds of phonological data (Lexical Phonology, Prosodic Phonology, co-phonologies, indexed constraints etc., with Distributed Morphology being a semi-exception).⁴

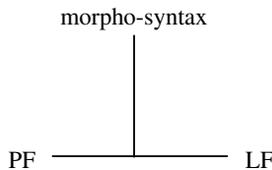


Figure 16.1 The inverted T model: All concatenation before all interpretation

In the perspective of the inverted T, (morpho-)syntax is the central concatenative system which has the privilege of concatenation: its output is sent to two interpretational modules, phonology (PF) and semantics (LF), which assign a form and a meaning to the morpho-syntactic structure.⁵ Morpho-syntax and the two interpretative modules are thus procedurally ordered so that words and sentences are pieced together before being shipped off to interpretation at PF and LF. That is, the function of phonology in the (production-oriented) generative perspective is to translate morpho-syntactic structure into a code that may be used by the sensory-motor system. In perception, the movement is reversed: phonology helps to interpret the phonetic signal, that is, to identify morphemes.

'Phon' in the word phonology is thus an (understandable) misunderstanding: natural language is modality-independent. That is, linguistic structure may run through any channel as long as humans are able to produce and perceive contrast. Two modalities are known to be used by humans in order to ship messages in natural language: vocal and signed. The former is the default, but the latter is used if the former is not available (there are documented cases where deaf children develop a sign language *ex nihilo*, that is, in absence of any linguistic stimulus, for example, Senghas et al., 2004). A more accurate definition of phonology and its function is thus the translation back and forth between a physical signal and morpho-syntactic structure: phonology is a translational device.

SPE also adds an extra proviso (which does not follow from the inverted T structure): all concatenation must be done before all interpretation. While the inverted T model remains predominant up to the present day, the full completion of concatenative activity before any interpretation can begin will be a bone of contention in further development. The competing view is so-called interactionism according to which concatenation and interpretation are interspersed. This is the key idea of Lexical Phonology (Kiparsky, 1982a and b; Hargus and Kaisse, 1993 provide an overview), and today also the backbone of minimalist syntax (derivation by phase, Uriagereka, 1999; Chomsky, 2000: *passim*).

A direct consequence of the modular architecture is the necessity for *translation*: every module works on its own proprietary vocabulary (this is how modules are defined in CogSci: domain-specificity, for example, Hirschfeld and Gelman, 1994). Morpho-syntax, for example, reacts on things like number, gender, animacy and so forth, which is what the morpho-syntactic part of lexical entries is made of. By contrast, labiality, occlusion and the like are entirely transparent to morpho-syntax. On the other hand, phonology works on these, but is unable to parse number, gender and the like.

Therefore, SPE and all subsequent generative theories establish a translational mechanism that transforms morpho-syntactic structure into objects that can be read by the phonological computation. In SPE, morpho-syntactic structure is translated into hash-marks # by a mapping algorithm (all major categories as well as projections thereof are preceded and followed by a #), later on into prosodic constituency (Prosodic Phonology, Selkirk, 1981b: *passim*). Even though on other grounds and following a different reasoning, structuralist practice was also modular in the sense that so-called Level Independence enforced translation between phonology and morpho-phonology (see Scheer, forthcoming a). The only thing that changes over time is thus the output of the translational process, which is defined in terms of the phonological units that were used in the respective theories: juncture *phonemes* in structuralist times, hash marks which were defined as a special kind of segment in SPE (#s were [-segment] segments, while 'real' segments were [+segment]), and finally

autosegmental constituency in the early 1980s when all areas of phonology were autosegmentalized.

On the other hand, it is also true that SPE violates the translation requirement since apart from hash marks, it also makes brackets and labels available in the phonology, that is cyclic structure (on which more shortly) and genuine morpho-syntactic vocabulary such as verb, noun, adjective. For example, a word such as *theatricality* appears in phonology as [[[theatr]_N ic + al]_A i + ty]_{NV} and phonological computation can freely refer to the information that this or that portion of the string is a noun, a verb or an adjective. Hence, typical SPE-style rules such as ‘k palatalizes before e, but only if this e is the dative singular’. We will see below that reference to so-called morphological diacritics is one of the reasons why SPE came under fire in the 1970s. The problem was only (partially) solved in the mid-1980s when Prosodic Phonology established the so-called Indirect Reference Principle (e.g. Nespor and Vogel, 1986) which prohibits reference to untranslated morpho-syntactic information in phonological processes.

2.2 Cyclic Derivation

Another key property of generative thinking, cyclic derivation, is directly related to the modular architecture. The idea that phonological (and semantic) interpretation follows morpho-syntactic structure, that is, proceeds from inside-out (from the most to the least embedded constituent) bears the stamp of a generative copyright. That is, a string with the morpho-syntactic structure [[[A] B] C] is computed in such a way that first A is interpreted alone; the result is then assessed together with B (i.e. AB), and finally the product of this computation is interpreted together with C, that is, ABC. It is important to note that the computational system, that is, the set of instructions that cause modifications, is invariable: whatever the string, it is interpreted by the same phonology. Also, there were no requirements regarding so-called strict cyclicity (a notion that we return to below) in SPE, and Chomsky and Halle (1968) are explicit on the fact that cyclic derivation concerns morphological as much as syntactic structure: it applies ‘to all surface structure whether internal or external to the word’ (Chomsky and Halle, 1968: 27).

Cyclic derivation represents a genuinely new idea that was absent from linguistic thinking in Antiquity and the Middle Ages, and from the neogrammarian and structuralist paradigms in modern times. It first appeared in Chomsky et al. (1956: 75), but was codified only in SPE, where it is called the Transformational Cycle. In the 1970s, it was also known as the Phonological Cycle (Mascaró, 1976), and today cyclic derivation (or cyclic spell-out) is known as derivation by phase, which is the spine of current minimalist syntax (Chomsky, 2000: passim).⁶

2.3 Anderson's Prism: Structure Versus Process

Finally, let us consider something that SPE does *not* have, and which will become central in the further evolution of the field: representations. The major insight of structuralism was the existence of the system and hence of systemic pressure: a sound is not a linguistic object per se – it acquires this quality only by virtue of its relationship with other sounds in a system. That is, the same physical object may have different phonological behaviour according to the system in which it is couched. This central piece of structuralist thinking was eclipsed in SPE (without argument).

Granting no role at all to structuralist structure, SPE certainly ranges close to the computational end of a scale whose extremes are computation and structure. It is true that SPE is not only made of computation, though: computation necessarily computes something – objects – which are thus different from computation. This is the dualistic philosophy of Cognitive Science: modules are input-output devices that take *objects* as input, transform them and return their altered versions.

What are thus the objects that are manipulated by computation in SPE? Feature-matrices that are made of binary features. While serialism (i.e. extrinsically ordered rules) is closely associated with generative phonology (albeit not its invention, see Bloomfield, 1939 and the discussion aroused by the (non-) reception of this article in generative quarters, as reported by Goldsmith, 2008), Chomsky and Halle have never claimed the paternity of binary features: these were introduced by Roman Jakobson (Jakobson, 1939a; Jakobson et al., 1952) in a structuralist environment. SPE has merely replaced their original acoustic definition by articulatory values. In this perspective, thus, extrinsically ordered rules change the values of articulatorily defined binary features, which are the only structure that SPE recognizes. Or rather, to be precise, SPE recognizes yet another type of item beyond binary features: segments, which are an aggregation of features. A segment is defined as a feature matrix whereby all segments are made of the same features and the same number of features. Segmental variation, then, is exclusively a matter of varying feature values.

The non-computational part of SPE, its structure, is thus made of features and feature-matrices. An interesting question that is not usually asked in the (contemporary or later) literature is whether it makes sense to talk of this kind of structure in terms of representations. As we will see below, the advent of autosegmental representations has also introduced the notion of *well-formedness*, which was central for phonological thinking in the 1980s. It is therefore reasonable, in any case useful, to establish a strong bond between representations and well-formedness: only representations can be well- or ill-formed, and something can only be a representation if it can be ill-formed. This, however, is clearly not the case with features and feature-matrices: there are some combinations

that either cannot or do not occur, but never because of an intrinsic impossibility that is defined by grammar.

For example, a vowel cannot be [+high] and [+low] at the same time – but this is for physiological reasons. And some language may well not possess this or that combination of feature values of the universal set of features – actually most of the astronomical number of logically possible combinations do not occur. But this again is for reasons that have got nothing to do with grammar: inventory selection and overgeneration will drive SPE into trouble as we will see below. Finally, core properties of features (such as their binary nature: they have either a positive or a negative value, no third choice is available) and of feature matrices (e.g. no feature may occur twice in a given matrix) are set in stone before the derivation begins: no lexical item can contravene, and no computational action can create monster features without value or monster matrices. By contrast, autosegmental representations may become ill-formed in the course of a derivation. They therefore introduce a new quality into phonological thinking, which in retrospect turns out to be a milestone in the history of phonology.

Let us now place this discussion in a larger context. The balance between structure and process is the prism that Stephen Anderson (Anderson, 1985) uses to understand the history of phonology in the twentieth century. Anderson has detected a regular see-saw movement between theories that stand far on one side of the spectrum, and others that approach the opposite extreme. Also, both properties are in a reverse proportional relationship: when one goes up, the other goes down.

One idea is that the balance between representation and computation is a valid instrument for understanding what the constant changes in terminology, theories, concepts and schools is all about. Writing at the representational peak of the 1980s, Anderson extrapolates that phonology stands at the dawn of a new computational, hence anti-representational era. Here is the last sentence of his book:

If current attention to the possibilities of novel sorts of representations leads to a climate in which the importance of explicit formulation of rule-governed regularities disappears from view, the depth of our knowledge of phonology will in all likelihood be poorer for it. We hope that this book has demonstrated that neither a theory of rules nor a theory of representations constitutes a theory of phonology by itself. (Anderson, 1985: 350)

Little did he know how right he was, that is, how far on the extreme computational end OT would take phonology a couple of years later. Given its predictive success, the structure versus process prism will be used below as a measure of

the evolution of the field (van der Hulst, 2004 also looks at the history of phonology through this lens). Anticipating the facts to be reported, we may note that the generative micro-cosmos (as compared to Anderson's larger span) has achieved a complete loop from an almost exclusively computational theory over the heavily representation-oriented autosegmental 1980s back to the exclusively computational environment of OT (where literally no structure is left since even structure is supposed to 'emerge', that is, to be the result of computation). A reasonable comment on this situation is certainly that reinventing the wheel every now and then, or going in circles, does not witness the maturity of a field.

3. Revolution and Revision

3.1 What is Wrong with SPE: Creating a Revolutionary Situation

3.1.1 *Overgeneration*

The fundamental problem of SPE is that it can describe all phonological processes that exist, as well as those that do not. For a theory that sets out to build a system (a grammar) that is able to generate all structures that are attested (or well formed), and none that are not attested (or ill formed), this situation amounts to a declaration of bankruptcy. Chomsky and Halle (1968) had lucidly spotted the problem in the ninth chapter of SPE, but the remedy that they suggested, a theory of markedness, never really emerged.

Overgeneration concerns both structure and computation, and is regularly pointed out in the post-SPE literature (e.g. Goyvaerts, 1981; Kaye, 1989: 58ff.). It was already mentioned that given free combinability of positive and negative values, the fixed (and universal) set of binary features (say, 20) creates a bewildering number of logically possible segments, of which only a vanishing fraction exists (either in any given language or cross-linguistically, even when physiological impossibilities are excluded).

On the computational side, the universal rule format $A \rightarrow B / K$ does not impose any restriction on what kind of object can instantiate A, B and K. Anything may be turned into any other thing (no restriction on the relationship between A and B): $n \rightarrow \eta / _ k$ is just as plausible as $n \rightarrow f / _ k$ or $r \rightarrow \int / _ k$. Also, anything can provoke any change (no restriction on the relationship between $A \rightarrow B$ and K): $n \rightarrow \eta / _ k$ is just as good as $n \rightarrow \eta / _ p$.

It is obvious for most phonologists⁷ that natural language does not work like that. Certain processes are recurrent (e.g. dental nasals become velar before a velar obstruent, dental nasals become labial before a labial obstruent), others are rare, and some do not occur (e.g. a [s] that is turned into [b], see Ewen and van der Hulst, 2001: 3ff. on this). It is therefore wrong to leave the variables A and B without any restriction.

It is also self-evident for most phonologists that any context may not provoke any effect: there are precise causalities, and it is not reasonable to assume that a given object may trigger any process and its reverse. Hence, the relationship between K and A is not free but obeys precise causal patterns (e.g. dental nasals may become velar before a velar, but not before a labial stop).⁸

Finally, wild overgeneration is also promoted by so-called morpho-phonology: as has already been mentioned phonology could freely refer to morpho-syntactic labels (e.g. 'velars palatalize before front vowels, but only if these represent the dative singular') since SPE granted free access to labelled brackets that were inherited from morpho-syntax.

3.1.2 Related Items Must Have a Common Underlying Form

Another area where overgeneration produced strange blooms is the relationship between etymologically, paradigmatically or semantically related items. In SPE and early generative phonology, it was admitted without discussion that items for which the linguist can determine a relatedness of this kind must have a common underlying form from which the surface variation is derived by phonological rules. Otherwise, it was argued, a generalization is missed.

For the sake of illustration, it is useful to choose the wildest case on record: the work by Theodore Lightner is certainly not representative of the average post-SPE phonologist, but the fact that the theory allowed Lightner to posit such derivations without hesitation is an indication that something was wrong. A (phonological) theory must somehow define what a possible derivational relationship is, and what is not.

Lightner (1978: 18f., 1981, 1985) holds that the following pairs are derived from a common underlying form: *eye* and *ocular*, *thunder* and *detonation*, *dental* and *tooth*, *rebel* and *bellicose*, *cardiac* and *heart*, *three* and *third*, *gynaecology* and *queen*, *sweet* and *hedonism* and so on. Since the alternations h-k (*heart* - *cardiac*), d-θ (*third* - *fourth*) and s-h (*sweet* - *hedonism*) suppose Grimm's Law, Verner's Law and the Ancient Greek s > h shift, Lightner concludes that these processes are performed by the grammar of present-day English speakers.

What these examples show is that the more distant (etymologically, paradigmatically, semantically) two items that are supposed to be related by a phonological derivation are, the more demanding they are for underlying representations and for phonological computation. That is, the underlying representation that needs to be set up for *sweet* and *hedonism* is highly *abstract*, that is, fairly distant from the surface form (of at least one of the items). In addition, the phonological rules needed to carry out the transformation are numerous and make reference to diachronic processes that occurred several hundred, sometimes several thousand years ago. Therefore, a typical feature of SPE-style analyses is to propose synchronically underlying forms that are close to or identical with diachronically distant stages of the language (in the case of

SPE Old English, sometimes Common Germanic), and to set up rules that mimic diachronic evolution. It is not quite plausible, though, that present-day natives have acquired Old English lexical forms and the diachronic rules that transformed them over centuries.

The tacit rule that SPE and post-SPE phonology obeyed was thus as follows: a bulky (phonological) computational component associated to a small number of lexical entries make a better world than the reverse situation, or anything in between for that matter. The origin, development and (non-)justification of this view is discussed, for example, in Foley (1977), Kenstowicz and Kisseberth (1977, 1979); Anderson (1985: 331f.) provides an informed overview.

The three alternatives to a phonological derivation that is based on a common underlying form are known and have always been practiced: related items such as *electri[k]* and *electri[s]-ity* can either represent two distinct lexical entries (*electricity* is stored as a single lexical entry, which means that its pronunciation does not involve any morphological concatenative or phonological activity).⁹ There may also be a case of allomorphy (there are two stems recorded in the lexicon, *electri[k]* and *electri[s]*, which are selected by a morphological, rather than a phonological context: the morpheme *-ity* selects the latter). Finally, the third possibility is suppletion: *good* and *better* are certainly two distinct lexical entries; as is the case for allomorphy, they are in complementary distribution, but the selection is made according to a grammatical, rather than a morphological context.

On this backdrop, then, the two issues discussed – overgeneration and derivational relatedness – defined the research agenda in the 1970s. To cut a long story short, everybody tried to reduce the expressive power of the grammar – its abstractness. By what means exactly this was to be done was the dominant question. In any event, all phonological theories that emerged in the early and mid-1980s have to some extent learned the lesson that many alternations which early generativists believed were produced by online phonological computation do not represent a synchronically active process. The set of alternations that phonological theory is called to account for, then, is far smaller than what SPE thought it was: phonology has to shrink in order to be viable. This insight had a revolutionary and a revisionist offspring.

3.2 Revolution: Natural (Generative) Phonology

The major challenger of SPE in the 1970s were the two Natural Phonologies, one generative, the other not. The initial spark of the latter is David Stampe's Ph.D (Stampe, 1972b), which evolved into Natural Phonology (NP) (Donegan and Stampe, 1978, 1979; Dressler, 1974, 1984; Hurch and Rhodes, 1996; Dziubalska-Kořaczyk, 2002). The former, Natural Generative Phonology (NGP),

was founded by Theo Vennemann (1971, 1976 [1971]¹⁰) and developed namely with his student Joan Hooper (today Bybee: Vennemann 1972a and b, 1974a and b; Hooper, 1975, 1976).

Both Natural Phonologies share a basic set of assumptions and principles, but as indicated by their name one is more revolutionary than the other: while NGP accepts the basic generative architecture, NP has left generative grounds, namely on the count of functionalism, which is endorsed.¹¹

The Natural Phonologies promote a radical means of reducing the expressive power of the grammar: they cut down the set of alternations that represent phonological activity by some 80% or 90% in comparison to SPE (my estimate). In order to do that, a structuralist notion is revived that was abandoned by SPE: the morpho-phonological level. That is, N(G)P does not deny that an alternation such as *electri[k] - electri[s]-ity* is due to online computation – it represents morpho-phonological, rather than phonological computation. In other words, we face allomorphy.

In NGP, the two criteria that divide alternations into one or the other type are the True Generalization Condition (Hooper, 1976: 13ff.) and the No-Ordering Condition (Hooper, 1976: 18ss.). According to the former, only phonetically accessible information can be used in the formulation of phonological rules, while the latter prohibits rule ordering. On this backdrop, alternations that do not suffer any exception in the entire language and exclusively appeal to phonetically retrievable information are called natural and granted phonological status (P-rules). By contrast, alternations that are not surface-true and/or whose statement requires non-phonetic, that is, morpho-syntactic information, are rejected into the pool of morpho-phonemic rules (MP-rules). This contrast is the NGP-adaptation of Stampe's (1972b) original distinction between processes (P-rules) and rules (MP-rules).

On Stampe's view, the former are natural, innate, productive, unsuppressible and effect minimal structural changes (hence they apply in loanword adaptation and interfere when non-native languages are spoken), while the latter are conventional, learned, dispensable (they do not interfere in loanwords and when foreign languages are spoken) and may be responsible for structural changes that involve more phonetic distance. In this environment, Stampe's perspective on language acquisition is that it involves the suppression of those innate processes that are not effective in the language at hand. As was mentioned, the split between processes/P-rules and rules/MP-rules reinstalls structuralist morpho-phonology, and hence Level Independence in its right.

This restrictive definition of phonology attacks two of the three sources of overgeneration in SPE: reference to morpho-syntactic information (which is prohibited for P-rules) and rule ordering (which is also ruled out in NGP). The third kind of overgeneration is the arbitrary relationship between the context of a rule and the change that it causes. Here it is hoped that the

restriction to phonetically accessible information will automatically introduce sound causality.

This setup was accompanied (or even partly provoked) by a diachronic reasoning regarding the life cycle of alternations. That is, what we see today in alternations such as *electri[k]* - *electri[s]-ity* are vestiges of a once-phonologically controlled alternation that has aged (e.g. Vennemann, 1972a). Alternations are born as phonetic regularities before moving into grammar where they are first phonological but at some point start to be riddled with morphological conditions, followed by lexical factors, and finally are levelled out or eliminated from the language by some other means. Therefore, asking the question how much of what we see is controlled by phonology is if not identical, at least concomitant with the question how much diachronics is in synchronic sound patterns.

In sum, then, NP is a radical answer to the overgeneration problem, and the label revolutionary is certainly adequate, be it only in recognition of the fact that extrinsic rule ordering, a prime candidate for the generative identity, is outlawed in NGP, and that NP is based on functional considerations. NP is thus anti-abstract in the sense that it restricts derivational depth (i.e. the distance between underlying and surface forms) to a minimum. As a consequence, phonology looks quite phonetic in its mirror: only phonetically defined items can be taken into account by phonological rules (P-rules). It needs to be noted that the three ingredients mentioned – prohibition of morpho-syntactic information in phonology, anti-abstractness and the reduction of phonology to phonetic and surface-true statements – are logically independent. Gussmann (2007), for example, endorses the former, but rejects the latter two.

3.3 Revision: Kiparsky

3.3.1 *How Abstract is Phonology?*

The revisionist research programme was laid out in Kiparsky (1968 [1973]) and realized through the 1970s mostly by Paul Kiparsky himself.¹² It attacks only one source of overgeneration, the one that is mentioned in the title of Kiparsky's seminal article: abstractness, that is, the 'derivational distance' between underlying and surface forms. The two other sources of overgeneration, reference to morpho-syntactic information and the arbitrary relationship between the context of a rule and the change that is caused, are left untouched.

Abstractness is attacked along two lines: Kiparsky restricts possible underlying forms, and he marshals the complexity of the derivation by two means: the restriction of computation to so-called derived environments and the requirement for rules to use material that is freshly introduced on the last cycle. The two restrictions on computation are united under the roof of Kiparsky's version of the Strict Cycle Condition (SCC).

The Alternation Condition (Kiparsky, 1968 [1973], 1973b) defines what a possible underlying representation is: in case a morpheme shows no alternation on the surface, it must not be any different in its underlying form. This results in a ban against so-called absolute neutralization, that is, items in underlying forms that never appear on the surface (on which more shortly). On the other hand, Kiparsky (1973b) restricts the application of rules in such a way that a certain rule class may only target derived environments. This was called the Revised Alternation Condition: obligatory neutralization rules apply only in derived environments. An environment is derived if it is created by the concatenation of two morphemes, or by the application of a phonological rule. This embodies the idea that phonological processes do not apply to monomorphemic strings, that is, when the trigger and the target belong to the same morpheme.

Trisyllabic Shortening (or Laxing) may illustrate the ban on absolute neutralization and the quest for derived environments. The process produces alternations whereby a long vowel or diphthong of bisyllabic items appears as a short vowel when a suffix is added: *div[aj]ne* - *div[ɪ]nity*, *op[ej]que* - *op[æ]city* etc.¹³ Monomorphemic items such as *n[aj]ghtingale* (*nightingale*) and *[aj]vory* (*ivory*), however, systematically resist the process, even though they satisfy the trisyllabic condition.

SPE reacts in a non-systematic way that misses the obvious morphological generalization and sets up abstract underlying forms: the application of the rule is eluded simply by destroying either the target or the triggering context of each individual lexical item. Instead of /aj/, *nightingale* is said to have an underlying /i/, that is, /nixtVngael/ (Chomsky and Halle, 1968: 234); and instead of /i/, the last segment of *ivory* is made a glide, that is, /ivorj/ (Chomsky and Halle, 1968:181). Independent rules that are ordered after Trisyllabic Shortening then take /ix/ and change it to [aj] (via /i/), and vocalize the final glide of *ivory*.

The Revised Alternation Condition kills two birds with one stone: (i) it dispenses with the absolute neutralization of /i/ in /nixtVngael/ and /j/ in /ivorj/ (which never appear on the surface) because it (ii) offers a different reason for the non-application of the rule, that is, the request for the triggering environment to be morphologically complex.

Another case where the underlying form of invariable morphemes could be distinct from their surface form are so-called free rides. In our example, these concern the converse surface situation, that is, cases where the third but last vowel of a monomorphemic item is short. Hence Kiparsky (1982b: 148) points out that on the SPE analysis, the underlying forms of [æ]libi, c[æ]mera or P[æ]mela cannot be determined: the third but last vowel could either be faithfully short and hence appear as such on the surface, or it could be underlyingly long, that is, /aj/libi, /kaj/mera, /paj/mela; in this case, the /aj/ will be turned into [æ] by Trisyllabic Shortening – a free ride, that is, one without consequences.

The ban on absolute neutralization also does away with this unwarranted indeterminacy.

Finally, the restriction of rule application to derived environments allows phonology to do away with some of the incriminated reference to morpho-syntactic information. For example, vowel shortening applies to *mean* [miin] – *meant* [ment] but not to *paint* [pejnt], *pint* [pajnt], *mount* [mawnt] because the latter are morphologically simplex, and their long vowel is specified as such in the lexicon (rather than derived by rule): /miin+t/ versus /pejnt/. Shortening, then, applies only to derived environments (Kiparsky, 1985: 87).

On the other hand, Kiparsky's attempt to define a reasonable line of division between computation and the lexicon also means that certain pieces of morpho-syntactically conditioned phonological processes need to be defended against the *concretist* ambition. By the late 1970s, the abstract SPE-mainstream was opposed by so-called concrete analyses that followed the principle 'the more concrete an analysis (i.e. the less distance between underlying and surface forms) the better it is' (e.g. Leben and Robinson, 1977; Tranel, 1981).

This line of thought struggled with the fact that all attempts at devising a formal measure of different degrees of abstractness failed: a so-called evaluation measure (or evaluation metric) could not be set up (Kiparsky, 1974; Campbell, 1981; Goyvaerts, 1981). In this context, abstract analyses with no other limitation than the learnability of rules (cf. Skousen, 1981) persisted (e.g. Dinnsen, 1980; Gussmann, 1980; Dresner, 1981). Kenstowicz and Kisseberth (1977: 1–62) offer extensive discussion of the issue. They argue in Kenstowicz and Kisseberth (1979: 204ff.) that in some cases, there is no alternative to the abstract option.

This is also the position of Kiparsky, who favours grammar-internal principles such as the Revised Alternation Condition and the SCC (to be discussed next) and is not prepared to do anything just in order to pay tribute to concreteness. Lexical Phonology, the theory that emerged from the abstractness debate in the early 1980s, may be viewed as an attempt at maintaining as much morpho-phonology as possible in the computational device of phonology while cutting away the wildest outgrowths of unrestricted SPE (see Scheer, forthcoming a).

Finally, Kiparsky merged the condition on derived environments with Chomsky's (1973) Strict Cycle Condition that was originally devised for syntax and later on applied to phonology (Kean, 1974; Mascaró, 1976). Chomsky's (1973: 243) original formulation is as follows: '[n]o rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node.' The effect is the blocking of rules whose structural description is met by a string which is made exclusively of material that belongs to previous cycles. That is, given $[[AB]_i C]_j$, a rule that is triggered by AB can apply at cycle *i*, but not at cycle *j*. Or, in other words, rules must use material that is introduced on the latest cycle, a restriction that prohibits multiple application of a rule.

Chomsky's (and Kean's and Mascaró's) condition on the applicability of rules is entirely irrelevant for derived environment effects: it will not prevent rules from applying to monomorphemic strings since these have necessarily been introduced on the latest (the only) cycle. Thus, Trisyllabic Shortening (*s[ej]ne - s[æ]n-ity*) will happily apply to *n[aj]tingale* and *[aj]voory* under Chomsky's SCC.

Nonetheless, Kiparsky (1982a and b) introduces his version of the SCC (1) as if it were just a restatement of Mascaró's:¹⁴

(1) *Strict Cycle Condition (SSC):*

- a. Cyclic rules apply only to derived representations
- b. Def.: A representation φ is derived w.r.t. rule R in cycle j iff φ meets the structural analysis of R by virtue of a combination of morphemes introduced in cycle j or the application of a phonological rule in cycle j.

Kiparsky (1982b: 153f.)

Kiparsky's attempt to kill two birds ('use new material!' and derived environment effects) with one stone (his scrambled SCC) was considered an important achievement in the 1980s, but then turned out to lead to a dead end: the derived environment condition being riddled with counterexamples, Kiparsky (1993) himself declares the bankruptcy of his version of the SCC 10 years later.

This does not mean, however, that the original Chomskyan baby in Kiparsky's blended SCC ought to be thrown out with the bathwater. The idea that 'old' strings which have already experienced interpretation are unavailable for further computation appears in the guise of the Free Element Condition regarding stress assignment (Prince 1985, Poser 1986, Steriade 1988) and structure preservation regarding syllabification (e.g. Steriade 1982:84ff, Harris 1993). Kaye (1992, 1995) makes it a general condition on computation, and today it is the heart of syntactic Phase Theory, where it is known as the Phase Impenetrability Condition (PIC) (Chomsky, 2000: *passim*).¹⁵

3.3.2 *The Evaluation Measure and Common Underlying Forms*

Following the footsteps of Kiparsky's revisionist programme which was only marginally concerned with the question of common underlying forms, a substantial body of literature tried to outlaw Lightner-type excesses without however jumping to the NGP conclusion that phonologically driven alternations must be surface-true and may not make reference to any morpho-syntactic information.

Objective and measurable criteria were sought that could be applied to any given alternation in order to decide whether it is the result of phonological computation on the basis of a common underlying form or not. Candidate criteria that were discussed tried to measure what is natural, simple, elegant, phonetically

plausible, psychologically real or typologically invariant (e.g. Hellberg, 1978; Koutsoudas, 1980; Dinnsen, 1980), but this was quite inconclusive.

To date, the question remains open whether an item (such as *electric-ity*) that looks morphologically complex is really considered as such by the grammatical system. All modern theories have somehow swung into a midfield position: it is not good or bad per se to have a big or a small lexicon, or to do little or a lot of phonological computation. Arguments must be made on a case-by-case basis: while some alternating items represent two independent lexical entries for sure, the online computation of others is beyond doubt; the swampy midfield, however, is large enough for still much debate to come. Carvalho (2002a: 134ff.) provides extensive discussion of this question (and Carvalho, 2004, especially considers the role of analogy as a serious player).

The discussion regarding common underlying forms also runs under the header of (anti-)lexicalism (treating *electric-ity* as a single lexical entry is the lexicalist position). The parallel between the oscillating evolution in phonology and syntax in this area is quite remarkable: after a decidedly lexicalist period in the second half of the 1970s and all through the 1980s (triggered in syntax by Chomsky's 1970 remarks on nominalization), the minimalist environment in syntax, and OT in phonology, today have moved back to the anti-lexicalist heydays of the 1960s (e.g. Newmeyer, 2005; Williams, 2007).

4. (Unintended) Counterrevolution: Autosegmental Representations

4.1 A New Player, Ill-Formedness, Opens New Horizons and Buries the Revolution

In the second half of the 1970s, the SPE-mainstream was drifting in unfriendly waters without real research programme and without any perspective of a significant evolution. At the same time, the revolution initiated by Natural (Generative) Phonology gained ground. Kiparsky's revisionist work offered some relief, but could not produce more than some kind of SPE-Light. It is then that two movements appeared out of the blue, that is, without being prepared, solicited or envisioned, which acted as a fountain of youth that made generative phonology take a new start on the grounds of renewed premises (see also van der Hulst, 2004 on this episode).

One was Lexical Phonology whose initial spark was Siegel's (1974) discovery of affix ordering, that is, a non-phonological phenomenon, whose interleaving with associated phonological phenomena allowed for the establishment of a new architecture of grammar (i.e. a new positioning of phonology with respect to morphology and syntax), and for the unification with Kiparsky's revisionist programme.¹⁶

The other, to be discussed in more detail below were autosegmental representations (also called non-linear, multi-linear). These brought an entirely new dimension into phonological theory and certainly deserve to be called, also retrospectively, the most significant scientific advancement that generative phonology has made since its inception. A good indication of this status is the fact that autosegmental representations are the least common denominator of all phonological theories, present and past (including OT, at least in word, on which there is more discussion below).

Autosegmental representations are groundbreaking because they introduce a new player that was absent in neogrammarian as much as in structuralist thinking: ill-formedness. It was mentioned in Section 2.3 that SPE has structure (the material on which computation works, that is, features and feature-matrices), but no representations: a representation is something that can be ill-formed for *grammar-internal* reasons. That is, simple concatenation of morphemes or phonological computation can cause an autosegmental representation to be ill-formed – while nothing can be wrong with features and feature-matrices when they are pieced together or computed.

The potential of ill-formedness as an overgeneration-killer that restricts the expressive power of grammar was rapidly identified, and phonologists put much hope into this new perspective after the straining experience of the abstractness debate. Also, the overgeneration-avoidance virtue of autosegmental representations is *intrinsic* since their mere existence, associated with the statement of the (universal or language-specific) conditions of well-formedness, dramatically cuts down possible results. Note that this does not mean that representations are a filter only on outputs: they restrict possible phonological expressions wherever they occur: in the lexicon, in intermediate or in surface forms. A violation of well-formedness may either lead to the crash of the derivation, or cause phonology to act in order to repair the offending property (Goldsmith's 1976a: 27 initial tone-based conception only built on repair strategies; this line of thought is embodied, for example, in Paradis, 1988 and subsequent work; van der Hulst, 2004 provides an overview of the constraint-and-repair approach).

Representations are certainly a more powerful overgeneration-inhibitor than Kiparsky's revisionist programme, where a phonological computation could crash for computation-internal reasons. It could now also crash because it either works on or produces an ill-formed structure. This gives new impetus to Stephen Anderson's perspective on the equilibrium between structure and process: autosegmental representations take the generative cursor away from the computational extreme of SPE into some midfield position.

The notion of ill-formedness is of course not the only virtue of autosegmental representations (more discussion on this below), and it goes without saying that autosegmental representations had not solved the overgeneration problem

as such (e.g. they had no effect on the ‘static’ overgeneration inherent in feature-matrices, see Section 3.1.1). They had, however, brought home the promise of overgeneration-inhibition, at least partly (Lowenstamm, 2006 provides a good overview of the credits).

The new autosegmental possibilities opened up new horizons, and after an initial period of timid proliferation in the late 1970s the entire field was turned upside down when everybody participated in the autosegmentalization of all areas of phonology. This emulation rapidly and completely eclipsed the revolutionary vigour of NGP, which all of a sudden appeared toothless and simply petered out (see Laks, 2006 on this decline). By contrast, natural non-generative phonology continued its development outside of the generative paradigm.

4.2 Genesis and Properties of Autosegmental Representations

It is reasonable to daresay that Kiparsky’s revisionist programme associated to Lexical Phonology would not have been able to capture the victory over the revolution alone. It is only autosegmental representations that could drive revolutionary concerns out of the field. Interestingly, this was completely unintended: autosegmental representations came into being as a problem-solving mechanism whose counterrevolutionary virtue and impact on overgeneration was only discovered as things unfolded.

Autosegmental representations quite independently emerged in the analysis of a number of different empirical problems; tone (Leben, 1973; Goldsmith, 1974; Williams, 1976), syllable structure (Fudge, 1969; Kahn, 1976), stress and rhythm (Liberman, 1975; Liberman and Prince, 1977), Semitic non-concatenative (templatic) morphology (McCarthy, 1979a) and vowel harmony (Clements, 1977) have played an important role. The emerging non-linear atmosphere was condensed into a general autosegmental theory by John Goldsmith’s (1976a) Ph.D, which is usually credited as the initial spark of autosegmentalism (see Anderson, 1985: 347ff.; Lowenstamm, 2006 for a survey). This is correct in the sense that the decisive property of autosegmental representations beyond their graphic and non-linear aspect is introduced in Goldsmith’s Ph.D: the idea that a structure may be ill-formed (Goldsmith, 1976a and b defines well-formedness conditions).

In careful historiographic work, John Goldsmith himself (see also Clements, 2000a) has also identified pre-generative inspiration for non-linear thinking. This is most obvious for the syllable, which is known at least since the neogrammarian school of the nineteenth century and widely acknowledged in structuralist work (e.g. Pike and Pike, 1947; Hockett, 1955: 51ff.; Pulgram, 1970 offers an overview) as well as in NGP (e.g. Hooper, 1972). In addition to this, Goldsmith (1979a) for example, (see also Goldsmith, 1976b; Goldsmith and

Laks, forthcoming) mentions Harris' (1944) long components, Bloch (1948) and the prosodic analysis of the Firthian school of the 1940s. While all of these precursors (in hindsight) heralded the non-linear idea by promoting this or that aspect of it, none of them held all pieces in hand, not to mention their constitution in terms of a formal system. This condensation into a frame that also provides a graphic identity to multi-linear structure (which was absent in the structuralist and the NGP conception of the syllable: in the latter syllable boundaries '\$' were merely inserted into the linear string) has only occurred in the second half of the 1970s in generative phonology (Halle and Vergnaud, 1980 provide an early overview of the then recent history).

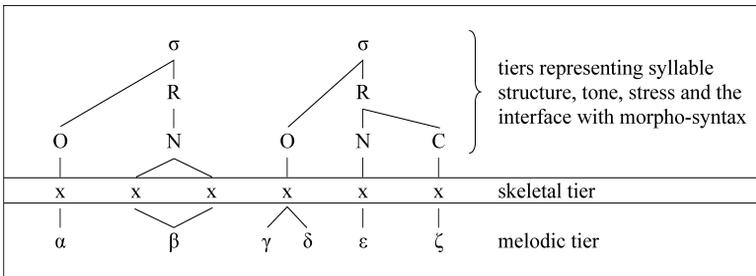


Figure 16.2 Shows a typical autosegmental representation

The idea of autosegmental representations is that two phonological objects (features in SPE) may be related in such a way that no linear order can be determined. This is why SPE is called a linear model (all objects, that is, features and feature-matrices, obey an unambiguous precedence relation with all other objects), against autosegmental representations, which are called non-linear. Under Figure 16.2, for example, there is no precedence relation between β and the two x-slots that it is attached to: β 'belongs' to one as much as to the other. Phonological objects may thus occur on different levels, which are called autosegmental tiers (or lines): in Figure 16.2 Greek letters belong to the melodic tier, x-slots to the skeletal tier, and higher items to the (or several) syllabic tier(s) (more on the specific identity of supra-skeletal tiers shortly). On every tier, items – which are called autosegments – obey a strict linear order. Non-linear effects are produced when items of different tiers are not associated one-to-one: association may be either one-to-many (e.g. β) or many-to-one (e.g. γ and δ).

This is the baseline that defines the minimal formal properties of autosegmental representations. Two more properties are agreed upon by all phonologists as far as I can see, although they do not follow from formal requirements. For one thing, autosegmental tiers are disposed in three-dimensional space, rather than only in two dimensions as in Figure 16.2. That is, the melodic tier in Figure 16.2 could well entertain associations with both the skeletal and the

syllabic tier. This specific configuration does not occur in practice, though, because of the second broad agreement, that is, the division of the autosegmental space into two basic areas, one accommodating melodic representations, the other constituent structure of various kinds: syllables, stress, interface with morpho-syntax (it is true that tone, which a priori is a piece of melody, commonly appears on the constituent side).

The red line that separates the melodic and the constituent area is the skeleton, which has a special status among autosegmental tiers. Skeletal slots are timing units that define the chronological progression of the overall structure: they are the instrument that defines the linear order of the non-linear structure. Items on other tiers may be related in various ways among themselves and with skeletal slots. When it comes to transform the structure into a linear phonetic signal, though, the skeleton is parsed slot by slot (left to right), and every autosegment of whatever tier that is attached to a given slot will be pronounced on that particular slot. This also means that items that are not attached to any skeletal slot (such as so-called floating segments, on which more is discussed below) are not pronounced.

Indeed, an important ground rule that is shared by all versions of autosegmental representations is that a result on the phonetic side supposes three things: (i) a piece of melody, (ii) a constituent that is itself integrated into the constituency above the skeleton and (iii) an association line that relates both items. In the absence of either, no audible trace is produced (e.g. Itô, 1986). This is illustrated in Figure 16.3, where the skeleton is used for the sake of exposition (in moraic theory, other constituents take over its role).

(1) autosegmental conditions on phonetic action			
a. phonetic	b. no phonetic	c. no phonetic	d. no phonetic
result: [α]	result	result	result
x		x	x
α	α		α

Figure 16.3 Autosegmental conditions on phonetic action

There are two autosegmental theories that in one way or another do not fit the picture presented thus far. Moraic theory (Hyman, 1985; Hayes, 1989a) is based on the claim that the skeleton does not exist. Phonologically relevant units are based on weight (morae), rather than on chronological precedence. Moraic theory thus challenges standard syllable structure, which is x-slot-based. It assigns morae to segments that contribute syllabic weight, that is, to vowels (and syllabic consonants) in all languages, parametrically to coda consonants in some, but never to onset consonants.

The other theory to be mentioned in this context is based on what is called the metrical grid (Prince, 1983; Hayes, 1984; Idsardi, 1992; see Halle, 1998 for an overview). Metrical grids are erected over segments and indicate their relative prosodic prominence, that is, regarding stress and rhythm. Prominence is marked by piling up a number of diacritic *grid marks* (which are graphically represented as '*'). Metrical grids thus compete with regular autosegmental arborescence (feet) for the representation of stress and, depending on the particular implementation, also with other autosegmental structure (e.g. the prosodic hierarchy that is responsible for the interface with morpho-syntax in Selkirk, 1984b). Metrical grids are clearly a child of autosegmentalism and the prevailing non-linear atmosphere of the 1980s; however, they are not autosegmental themselves because they are in no way non-linear: every segment has its own pile of grid marks, which may shift from one segment to another through computation, but which do not interrupt the linear sequence of segments. That is, a given grid mark is never related to more than one segment, and a given segment never associates to a grid mark that dominates a neighbouring segment. I leave the autosegmental status of metrical grids an open question here.

Let us now consider ill-formedness, the inherent overgeneration-inhibiting virtue of autosegmental representations. A very early and intuitive condition on well-formedness that remains in place to date is the ban on line-crossing (Goldsmith, 1976a and b): two items of the same tier (not of different tiers, though, recall that the autosegmental space is three-dimensional) may not entertain associations with another tier if they result in the crossing of association lines. In Figure 16.3, for example, /ε/ could not be associated to the x-slot of the coda, and /ζ/ simultaneously to the x-slot of the nucleus. The autosegmental literature has produced a number of well-formedness conditions, which have been debated over the years. Three prominent cases in point are the Obligatory Contour Principle (OCP), the Strict Layer Hypothesis (SLH) and the ban on two empty nuclei in a row. The OCP (Goldsmith, 1980; McCarthy, 1986; Yip, 1988) prohibits two identical (or similar) autosegments in a row on the same tier. It has played the role of an output filter for derivations, and also of a trigger for various repair strategies. The SLH concerns the constituency of the Prosodic Hierarchy (Selkirk, 1981b; Selkirk, 1984b: 26f.). It expresses the idea that a prosodic constituent of a given layer only dominates constituents of the immediately lower level, and is exhaustively contained in a constituent of the immediately higher level. Hence, there can be no nested constituents, and no association line can bypass a layer.

Finally, a ground rule in Government Phonology (Kaye et al., 1990; Harris, 1994) is that a structure which accommodates two empty nuclei in a row (in the same domain) is ill-formed.¹⁷ Hence, in French *il faut y revenir* 'one needs come back to this' where the two e's of *revenir* are schwas, either can be left out (*r'venir* is ok as much as *rev'nir*), but not both (**r'v'nir* is agrammatical). On the analysis

of Government Phonology, this is because the non-pronunciation of a vowel only means that the association between the relevant piece of melody and the nucleus is interrupted. The nucleus itself is still present and thus empty, and two empty nuclei in a row as in **r'o'nir* are ill-formed.

4.3 Autosegmentalization of All Areas of Phonology, Blooming Theoretical Diversity

The potential of autosegmental representations, regarding overgeneration but also a host of analytic advances, was well understood in the early 1980s. Therefore, early spartan representations were continuously enriched throughout the 1980s, leading to rather complex structures. This evolution was parallel to the expansion of arboreal structure in (GB) syntax during the same period.

In contrast to the theoretical monoculture of the 1990s and 2000s (the large array of OT-internal variation notwithstanding), a blooming variety of new theories emerged in the early and mid-1980s on the autosegmental ticket. As far as I can see, this is the first time in the history of generative phonology that the mainstream divided into clearly distinguished, partly or completely competing theories that were defined by original assumptions and carried a cleaving name. The only clear precedent is Natural (Generative) Phonology – otherwise phonologists gathered around general tendencies or issues such as abstractness/concreteness.

A good witness of this extraordinary vitality are the first two volumes of the then newly established journal *Phonology Yearbook* (edited by Colin Ewen and John Anderson, today *Phonology*), which were published in 1984/1985 and became the voice of the autosegmental project (the two volumes edited by van der Hulst and Smith, 1982, were also a focal point). The *Phonology Yearbook* 1/2 contains a number of contributions that give Lexical Phonology a more precise shape after Kiparsky's (1982a and b) founding papers (Paul Kiparsky, Jerzy Rubach and Geert Booij, Jerzy Rubach, K.P. Mohanan, Patricia Shaw, including also the first overview by Ellen Kaisse and Patricia Shaw), and they herald the new thematic focus that dominated the 1980s: the internal structure of segments.

The founding statements of the new autosegmental theories of segmental structure also appeared in the two volumes of *Phonology Yearbook*. On the one hand, Feature Geometry (Clements, 1985) proposed an autosegmental arrangement of a set of articulatorily defined binary features that is more or less the same as in SPE. Each feature resides on an autosegmental tier of its own, and features are bundled under labelled nodes that define natural classes of segments, a new and important notion (that was condensed a year later in Elizabeth Sagey's 1986 Ph.D).

The alternative system of melodic representation rejects binary primes and instead favours so-called privative (or monovalent, holistic) melodic primes that (i) define items which are bigger than a single feature (e.g. the prime |l| represents the high front tongue body position) and (ii) produce contrast by being either present or absent (rather than by being always present but having two distinct values). The privative idea was not new: it was first aired by Anderson and Jones (1974), but now received three distinct implementations in Dependency Phonology (Anderson et al., 1985, condensed in Anderson and Ewen, 1987), Government Phonology (Kaye et al., 1985) and Particle Phonology (Schane, 1984) (see Botma et al., this volume, on features).

It is to be noted that the privative option has rapidly gained ground: it entered Feature Geometry under the header of underspecification (Archangeli, 1988; Pulleyblank, 1988). The idea is that plus-minus specifications of some (but not of all) features are absent from lexical representations and come into being through spreading from the context, or by feature-filling default rules that apply at the end of the derivation. Today the question is not so much whether there should be underspecification, but rather which features should be underspecified (see Hall, 2007 for an overview).

This implementation of the privative idea is still quite different from the aforementioned theories that are direct heirs of Anderson and Jones (1974), though: in underspecification approaches, features are not just temporarily unspecified for a value: rather, they are missing altogether. Melodic primes are different in nature (bigger than a classical feature), and they are independently pronounceable, which is not the case of (underspecified) features. Finally, underspecified features often afford a three-way opposition (plus, minus, unspecified), while truly privative systems only oppose the presence to the absence of a prime.

Another important issue was the kind of relationship that melodic primes entertain: while everybody agreed that relations between linguistic objects, in phonology and elsewhere, are asymmetric and hierarchical, opinions diverged as to how this insight should exactly be implemented. The most obvious take is the one of Feature Geometry, which simply extends autosegmental trees to the infrasegmental area. The alternative is the heart of Dependency Phonology, that is, a dependency relation: the idea is that one item is more important than the other(s) and contributes more to the result of the reunion (cf. also Liberman and Pince's 1977 labelled trees where sisters are either weak or strong). Government Phonology follows this line of thought (here the distinction is between a head and (an) operator(s)), while Particle Phonology expresses asymmetry by the number of copies of the same prime that contributes to the definition of a segment: [i] consists of the prime |l|, [e] of |l+a|, [ɛ] of |l+a+a| and [æ] of |l+a+a+a| (the presence of several copies of the same prime does not have any

impact on the result in other theories). Van der Hulst and Ritter (1999, 2003) provide an informed overview of the various implementations of the dependency programme.

Melodic representation of course was not the only sub-field that was autosegmentalised: all areas of phonology were recast in autosegmental terms. Figure 16.2 shows syllable structure above the skeleton, but this is only an arbitrarily chosen example. The representation of stress and the interface with morpho-syntax was also autosegmentalized: while the former is expressed in terms of feet since Libermann and Prince (1977) (with the aforementioned competing metrical grid), Prosodic Phonology has replaced linear carriers of morpho-syntactic information in phonology, that is, SPE-type boundaries #, with a multi-layered arboreal structure, the Prosodic Hierarchy (Selkirk, 1981b, 1984b; Nespor and Vogel, 1986; see Scheer, 2008, forthcoming a on this move). Prosodic Phonology also proposed a global unification of all types of constituent structure above the skeleton: the Prosodic Hierarchy encompasses (with some variation regarding, for example, morae and the clitic group) the following units of growing size in a single hierarchical structure (which is marshalled by the Strict Layer Hypothesis): mora < syllable (σ) < prosodic (phonological) word (ω , about the size of a word), prosodic (phonological) phrase (φ , about the size of an NP or a VP), intonational phrase (IP, an intonational unit), phonological utterance (U, about the size of a sentence).

Finally, it must be noted that in parallel to the autosegmentalization of phonology, the new research programme for generative linguistics that was laid out in Chomsky's (1981) Pisa lectures and gave rise to the Government and Binding framework in syntax very rapidly seeped through into phonology. Grammar was divided into principles and parameters: the ban on line-crossing, for example, was a (universal) principle, while there was a fair amount of parametric variation in the application of the OCP. The import of the new syntactic model into phonology as a major change in perspective is especially (and explicitly) recognized not only in Government Phonology (Kaye and Lowenstamm, 1984: 123, Kaye et al., 1985: 305, Lowenstamm and Kaye, 1986: 97f., 124ff.), as well as in Declarative Phonology (Scobbie, 1991: 7).

4.4 What about Computation?

In this rush for representational expansion, a fair question to ask is what happened to computation. At first sight, the answer is: nothing at all. All through the 1980s, computation was carried out exactly as before, only that the set of ordered rules now applied to representations, rather than to linear strings of segments and feature bundles. Computation was called to do labour that it did not do before, though, since the most widespread idea regarding

representations above the skeleton was that they are absent from the lexicon and hence need to be created by rule. This means that underlying representations were not really different from what they were in SPE (and from what they have returned to today in OT): a linear sequence of segments (except of course that the internal structure of segments was now autosegmental).

Syllable structure, for example, was erected over unsyllabified lexical entries by a syllabification algorithm on the grounds of major category information provided by the segments. In further evolution, bits and pieces of constituent structure were stored in the lexicon as well. Rubach (1986), for example, carries out the autosegmentalization of vowel-zero alternations in Slavic languages (where alternating vowels are called yers), whose major advance was to distinguish alternating (Polish *pies* - *ps-a* 'dog, Nsg, Gsg') from stable (*bies* - *bies-a* 'devil, Nsg, Gsg') vowels that are phonetically and phonemically identical by associating the latter, but not the former to the skeleton in the lexicon.¹⁸ On this analysis, thus, alternating vowels are floating, while stable vowels are lexically associated to the skeleton. The very notion of *floating segment*, which has rapidly gained ground, actually supposes the existence of some constituent structure in the lexicon: it only makes sense to talk about something that floats if there are other items that do not float, that is, are associated.

Another typical case of floating segments that require the presence of constituent structure in the lexicon is French liaison, where word-final consonants are pronounced or not according to whether the following word begins with a consonant (*les [le] cafés* 'the coffees') or a vowel (*les [lez] enfants* 'the children'). The standard autosegmental analysis of liaison is in terms of floating consonants (Encrevé, 1983, 1988). Government Phonology (Kaye et al., 1990) has gone farthest in this direction by assuming fully syllabified lexical entries.

Finally, the nature of computation was significantly changed by the fact that, operating over representational material, all it could do was to link or to delink autosegmental items. Indeed, the typical computational operation in an autosegmental environment is to *spread* some item, that is, to create an association line between it and some other unit. The palatalization of velars before front vowels, for example, is understood as the spreading of some palatal prime from the trigger into the target. This contrasts with pre-autosegmental rule application where the structural change operated the modification of a feature value. Therefore, Harris (1994: 111), for example, holds that the 'class of possible phonological processes is restricted to operations of delinking or spreading'.

We will see in the next section that computation, which was by and large ignored in the autosegmental commotion, will come back like a boomerang in the 1990s in order to set back the cursor on Anderson's structure versus process-scale right to where it stood in the 1960s (or even farther towards the computational extreme). Recall from Section 2.3 that Anderson (1985) predicted at the representational peak of the 1980s or rather, on the way uphill, that

computation will take its revenge: his outlook was that phonology stood at the dawn of a new computational, that is, anti-representational round.

5. Second Revolution: Anti-Derivationalism

5.1 Anti-Derivationalism without Argument

The second revolution rages against derivationalism in general and ordered rules in particular. It is far more serious than the first revolution for two reasons: (i) it attacks one of the deepest layers of generative thinking and more generally of Cognitive Science and; (ii) it combines an internal concern, that is, the growing discomfort with ordered rules, with an external solution, Parallel Distributed Processing (PDP) (i.e. parallel computation), which is the spearhead of connectionism. Connectionism, however, is the neo-empiricist (neo-behaviourist) alternative that started to challenge the rationalist standard theory of Cognitive Science in the second half of the 1980s (Rumelhart et al., 1986: *passim*). Section 5.5 discusses the connectionist import at greater length.

But let us start by looking at what anti-derivationalism actually is, and what its motivation was. Rooted in the properties of the universal Turing machine (Turing, 1936–1937), derivationalism lies at the heart of the standard theory of Cognitive Science that emerged in the 1950s, and whose application to linguistics produced generative grammar. It is the idea that computation in the mind involves a set of instructions that act on the input in such a way that it experiences step-by-step modifications which occur in a chronological and logical order where the output of step $n-1$ is the input to step n . Serialism (perhaps more appropriate a word than derivationalism), then, boils down to the existence of a set of extrinsically ordered instructions that produce a chronological and logical order of events (and hence of the action of computation).

In generative grammar, serialism incarnates as extrinsically ordered rules in phonology, and in early syntax as extrinsically ordered transformations. These were abandoned in the early 1980s when GB replaced them by so-called move α , a system where movement (computation) is free in itself, but marshalled by constraints on representations (e.g. Newmeyer, 1986: 163ff.). Move α represents an important turn in syntactic theory away from restrictions on computation itself (Chomsky's 1973 original Strict Cycle Condition, extrinsically ordered rules) in favour of a central role of well-formedness constraints on representations such as barriers, the Empty Category Principle (ECP), case checking and so forth. The autosegmental evolution in phonology that was described in Section 4.2 follows the same track: representations are marshalled by well-formedness conditions such as the OCP. On the first page of their book, Prince and Smolensky (2004 [1993]) explicitly draw on the evolution in syntax

and declare that their new theory extrapolates the timid phonological precedent into a formal system:

[o]ur goal is to develop and explore a theory of the way that representational well-formedness determines the assignment of grammatical structure. [. . .] The basic idea we will explore is that Universal Grammar consists largely of a set of constraints on representational well-formedness' Prince and Smolensky. (2004 [1993]: 1f.)

There is thus a causal chain beginning with the emergence of autosegmental representations, which produced the need for well-formedness conditions on outputs, which in turn were generalized to constraint-based computation.¹⁹ In practice, Prosodic Morphology (McCarthy and Prince, 1986), which fully explores the autosegmental tool, was instrumental as a precursor of constraint-based computation. It developed at least two central devices of OT: correspondence theory and alignment of prosodic and morphological constituents (the foreword to the 2001 edition of the manuscript, McCarthy and Prince, 2001, explains this evolution in greater detail). We will see below that for reasons to be determined, the implementation of the constraint-based programme in OT has initiated the dissolution of representations as an autonomous actor in grammar.

For the time being, it is enough to take stock of the fact that while generative syntax abandoned serialism since 1981, the representational blossoming of the early 1980s left serialism entirely untouched in phonology.²⁰ In the second half of the 1980s, though, a diffuse and inscrutable discomfort with ordered rules arose, which quickly turned into a vigorous and lasting antipathy. I have read through the literature of that period in search for indications *why* serialism is supposed to be wrong, or why it aroused scepticism – without success. I have also asked phonologists who have lived through that period: all confirm that there was indeed a deeply rooted antipathy against ordered rules, and that this feeling was shared by about everybody across theories, but that it somehow remained below the surface. The broad reference to the evolution in syntax set aside, nobody could name, and I was unable to hunt down, sources in print that would have explained why serialism is wrong, and why phonology needed to engage in a major cultural break.

I happen to be aware of the reason why Government Phonology participated in the anti-serialist movement: because Jonathan Kaye and Jean Lowenstamm considered extrinsic rule ordering empirically vacuous. According to them, examples where serial ordering of instructions is alleged to be critical are either based on erroneous data, involve misanalysis or concern processes whose properties disqualify them as instances of phonological computation (no plausible relationship between trigger and effect, exceptions, appeal to morphological

information). An example for erroneous data is Martin Joos' famous dialect B of Canadian English for which there is no evidence (Kaye, 1990c, 2008), but which was used by Bromberger and Halle (1989) as the litmus test for rule ordering (see note 21). Examples for processes that are not phonological in nature are Trisyllabic Shortening (or other traces of the great vowel shift), and velar softening (see the aforementioned *electri[k] - electri[s]-ity*). The trouble is that apart from Kaye (1990c) which is only a short notice about the non-existence of dialect B, there is no trace of this programme in print (Lowenstamm and Kaye, 1986: 97 mentions that the model of the authors is referred to as the 'no-rule approach', but does not say why).

In contrast to the non-overt sources of anti-serialism, the origin of parallel computation is evident: connectionism. Since Rumelhart et al., (1986), the central argument in favour of parallel computation was clearly made and pedagogically repeated (e.g. Rumelhart, 1989: 134ff.): the implementation of serialism in a neuron-based environment appears to be unrealistic given the computational complexity that would be required and the time that it would take to carry out all intermediate steps one by one. We know that neurons are not serially ordered in the brain, but rather multiply interconnected. Hence, in order to get to grips with a realistic implementation in the brain, several things must be done simultaneously, just like many neurons fire at the same time and thus transmit information simultaneously. This is why we need 'brain-style computation' (Rumelhart, 1989).

We thus have an argument that calls for the modification of grammar, which is a model of competence, under the pressure of performance, that is, neural implementation. It is interesting to observe that Prince and Smolensky (2004 [1993]: 215f.) categorically reject this reasoning ('[i]t is not incumbent upon grammar to compute'), which they consider a category mistake. The rejection of performance-based reasoning is their general-purpose shield against the numerous objections that were raised against the astronomical inflation of computational complexity in OT (see Section 5.4).

It is idle to further speculate where the general antipathy against serialism came from, how it spread in absence of written record, and how an entire field could throw overboard a fundamental piece of its identity without any discussion of the reasons. In his article *In defense of serialism*, Clements (2000b: 193f.) makes the obvious point that the rise of connectionism on the Cognitive Science scene has played an important role in the development of anti-derivationalism in phonology. It remains to be seen, though, whether the development in Cognitive Science really seeped through into phonological quarters at this early stage: it was only unfolding when anti-serialism was already widespread in phonology.

In any event, the fact is that the defenders of serialism – of which Morris Halle is the most prominent figure whose position has not varied since the

1950s – reacted on the anti-derivational atmosphere by exposing arguments in favour of ordered rules. The article by Sylvain Bromberger and Morris Halle published in 1989 (Bromberger and Halle, 1989) discusses the question whether the abandoning of ordered instructions by GB syntax and its replacement by the Principles and Parameters approach should lead phonology to follow the same track. The authors reject this perspective because, as they try to show,²¹ the subject matter of phonology and syntax is intrinsically different. Bromberger and Halle thus defend serialism against an extra-phonological trend, and also mention the fact that extrinsically ordered rules have come under fire in phonology. Significantly, though, they can come up with only two anti-derivational references: Lowenstamm and Kaye (1986) and a 1987 LSA (Linguistic Society of America) presentation by B. Majdi and David Michaels. This is good indication of the fact that the anti-derivational atmosphere was by and large absent from print: Bromberger and Halle were fighting against an invisible enemy.

5.2 Consequences of the Latent Anti-Derivationalism: New Constraint-Based Theories in the Early 1990s

The latent anti-derivationalism of the late 1980s was the driving force of the events in the early 1990s. Anti-derivationalism was now made explicit as the driving force behind the emergence of new theories, but there was still no discussion of comparative merits with serialism, or of the reasons why serialism is a bad thing to have. That is, parallel computation is typically introduced as an alternative to ordered rules, but authors leave it at this juxtaposition (e.g. Prince and Smolensky, 2004 [1993]).

Three theories that have emerged in the early 1990s (or late 1980s) are based on the anti-derivational mantra: Optimality Theory (Prince and Smolensky, 2004 [1993]), Declarative Phonology (Scobbie, 1991; Scobbie et al., 1996) and Government Phonology (Kaye et al., 1990, which was prepared by work quoted earlier since the mid-1880s). The computation in all three cases is based on constraints, which however do not have the same status: while they are ranked and violable in OT, they are absolute (i.e. non-violable) in Declarative Phonology (see discussion of constraints in Uffmann, this volume). Computation in Government Phonology was not explicitly regulated for some time (the only indication that could be found in print was Kaye's 1992: 141 statement according to which processes 'apply whenever the conditions that trigger them are satisfied'), but its constraint-based character is obvious since the mid-1990s (Licensing Constraints, Charette and Göksel, 1994, 1996; Kaye, 2001; Gussmann's 2007 recent book on Polish is an application to an entire language).

Constraints in Government Phonology thus apply whenever a form may be modified by them, but with no extrinsic ranking or ordering, and without being

able to be violated: the set of constraints (the ϕ -function in Kaye's 1995 terms) is (simultaneously and) iteratively applied to the string that is submitted to interpretation, and computation ends when no further modification can be made (this is an obvious parallel with Harmonic Serialism, on which more discussion below). To use serial vocabulary, this system is thus able to handle a feeding relationship (the input for the application of a constraint is created by the modification of the input string by another constraint), but no other (i.e. bleeding, counter-feeding, counter-bleeding). A difference must therefore be made between serial computation as in GP and serialism as such. In the former, computation is serial in the sense that constraints may apply to the same string several times, and that intermediate steps may exist, whereas in the latter there is no extrinsic or logical ordering of instructions, that is, classical extrinsic rule ordering. There is also no ranking or prominence relationship among constraints: all instructions are equally important, and there is no selective application (all items of the ϕ -function apply when a string is computed). However, Gussmann's (2007) book on Polish (as well as Polgárdi, 1999) indicates that constraints may also be ranked, that is, that in case of a conflict between two constraints one will be given priority (see Scheer, 2010 b).

Declarative Phonology is directly affiliated with HPSG (Head-Driven Phrase Structure Grammars) on the syntactic side and therefore does not stand on generative grounds: under the pressure of overgeneration, HPSG has taken the radical step to eliminate computation altogether. The result are so-called monostratal representations, which are fully informed with morpho-syntactic, semantic and phonological information that is available at any point in the derivation (talking about a derivation is actually improper because monostratalism rejects the existence of distinct underlying and surface representations). The major issue that HPSG has with the generative approach is thus about modularity: there are simply no modules in the HPSG landscape, which is a fully scrambled everything-is-one environment.

Regarding OT, it is to be noted that the juxtaposition of serial and parallel computation was really a matter of indecision in the early days: Prince and Smolensky (2004 [1993]) consider so-called Harmonic Serialism all through their manuscript. Harmonic Serialism works like regular OT, only that the candidate set produced by GEN is much more local, and the winner of the strictly parallel evaluation procedure is fed back into GEN. This procedure is repeated until no harmonic improvement can be achieved anymore. Hence, Prince and Smolensky (2004 [1993]: 6) write that

[d]efinitive adjudication between parallel and serial conceptions, not to mention hybrids of various kinds, is a challenge of considerable subtlety, as

indeed the debate over the necessity of serial Move- α illustrates plentifully [. . .], and the matter can be sensibly addressed only after much well-founded analytical work and theoretical exploration.

A more direct application of connectionism to phonology, John Goldsmith's Harmonic Phonology (Goldsmith, 1992, 1993; Larson, 1992), also follows this track: the successive application of rules progressively increases harmony, and the well-formedness of representations is measured in gradient, rather than in categorical terms. Finally, Harmonic Grammar (Legendre et al., 1990b; Smolensky and Legendre, 2006) is also a direct application of connectionism, in fact one which is most closely related to the interests of Cognitive Science: it is based on so-called weighted constraints, which like Goldsmith's gradual well-formedness are a direct transcription of the central connectionist notion of connection weight (and the activation level of neurons which defines their output). In this perspective, the relationship between constraints is one of lesser or greater prominence, rather than of strict dominance: less important constraints can league together and outrank a more important constraint on account of their cumulated weight.

Serialism, which stood unchallenged and without alternative since the 1950s, has thus first fallen into disgrace for reasons that are not clear and have not been made explicit, and was then replaced by parallel computation when this option was available as an import from connectionism (more on this import in the following section).

Parallel computation has an important corollary: it can only be done on the basis of constraints. We have seen in Section 5.1 that autosegmental representations and well-formedness conditions were the initial spark of the development of constraint-based computation. Evicting rules as the basic carrier of computational instructions in favour of constraints was thus the result of the conjoint action of parallel computation (which in turn was the instrument of anti-derivationalism) and the increasing importance of well-formedness conditions.

Another factor, whose impact on the movement towards constraints is difficult to evaluate, though, is what was known as global conditions and conspiracy in the 1970s: the latter notion is known at least since Kisseberth (1970a). It builds on the observation that a number of seemingly independent processes may conspire to produce the same result (such as the avoidance of closed syllables). In rule-based computation, each process requires its own rule, and the obvious generalization regarding the restriction on the output is missed. The relevant generalization may be captured by a single constraint on the output, and this is why current parallel computation is often called output-oriented.

5.3 How Computation Works: Rules Versus Constraints, Eventual Hybrid Solutions

Rules and constraints have different properties:²² rules are made of a structural change (the part on the lefthand side of the slash in $A \rightarrow B / C_D$) and a structural description (the part on the right-hand side). The latter defines a string, CAD, that is in need of modification. Constraints do the same thing (if in different vocabulary, see below): they issue general requirements or prohibitions, in our case, for example, *CAD. However, they do not specify what should be done in order to satisfy the requirement or prohibition at hand. This is another reason why constraint-based computation is said to be output-oriented: constraints specify how things should or must not look like, but do not give any indication how the desired or prohibited state of affairs should be achieved. That is, constraints divorce the structural description and the structural change of rewrite rules. Goldsmith and Laks (forthcoming) point out that this split was already suggested by Sommerstein (1974).

Hale and Reiss (2008: 195ff.) discuss the formal difference between rules and constraints at length. They describe rules as a function that maps an input representation to an output representation (i.e. something is modified), while a constraint maps an input representation to the binary set 'violation' or 'no violation', that is, without modifying anything. The equivalent of the modification that is operated by rules is achieved by comparing the input form to candidates, and by determining which candidate incurs the least harmful set of violations. OT-style parallel computation thus in fact falls into three independent steps which are serially ordered: (i) GEN: the candidate set is generated on the basis of a given input; (ii) EVAL: candidates are evaluated for violations of the constraint set. This is the time-saving (in the connectionist sense) and truly parallel piece of the computation since the evaluation of a candidate by different constraints can be done simultaneously: a given computational action does not need to 'wait' for another computational action since constraints assess candidates independently. (iii) Finally, the optimal candidate is determined once the results of the computational action of all constraints are known.

Hale and Reiss (1998: 196) build on this difference between rules and constraints in order to make a point against the latter from the logical and cognitive point of view: a grammar ought not to contain explicit statements against monsters (they use the NoBANANA example in order to show that there is no point in explicitly excluding real bananas from UG by an explicit statement therein). Constraints, however, only inform some other part of the grammar that a given representation is ill-formed, and there is an infinite number of ways in which a representation can be ill-formed. Hence, explicit statements (constraints of the NoBANANA kind) are needed in order to rule out the monsters.

Another important difference between constraints and rules is the vocabulary in which they are stated: while rules can only refer to the specific vocabulary items that phonology is made of (features or other items of autosegmental representations), constraints are made of prose statements and can express anything that prose can express (including very broad instructions such as 'be lazy!', which is the formulation of the constraint *LAZY* that Kirchner, 1998 believes is the motor for lenition; **STRUCTURE* is another case in point). This loss of reference to a specific phonological vocabulary is meaningful in terms of Cognitive Science: we will see below that while so-called domain-specificity is a defining property of cognitive modules (which thus operate over a specific and proprietary vocabulary), it is denied by connectionism, where computation is content-free.

The fact that the formulation of constraints is not constrained in any way is a well known and oft-mentioned property of this carrier of computational instruction (see, for example, Carvalho's 2002b formulation 'from constraint-based theories to theory-based constraints'), which to date however has failed to produce any effect: as far as I can see, no attempt is made in order to define what an (im-)possible formulation of constraints looks like.²³

Obviously, the freedom to express anything that prose can express dramatically increases overgeneration, which takes us back to the post-SPE debate. OT has often been charged with computational irresponsibility: in principle, GEN produces an infinite set of candidates that cannot even be stored, let alone computed, and the number of grammars that a set of, say, two or three hundred universal constraints (which is an extremely conservative estimate) produces given free ranking is astronomical (even when logically impossible rankings and those that produce identical patterns are counted out). Factorial typology is commonly advertised as a trump for modelling dialectal variation, but the hundreds, thousands or more systems that are produced without empirical echo are usually not mentioned. Also, the concern for limiting the number of constraints that was present in early OT has more or less disappeared from the agenda: new constraints are proposed every day, and hardly anybody is able to establish a comprehensive inventory.

From the beginning, the answer of OT to criticisms regarding excessive computational complexity (and overgeneration) was to call on the competence-performance distinction (but recall from Section 5.1 that it is precisely performance-based arguments that are used by connectionism in order to promote parallel, that is, brain-style computation):

It is not incumbent upon a grammar to compute, as Chomsky has emphasized repeatedly over the years. A grammar is a function that assigns structural descriptions to sentences; what matters formally is that the function is well-defined. [. . .] Grammatical theorists are free to

contemplate any kind of formal device in pursuit of these goals; indeed, they *must* allow themselves to range freely if there is to be any hope of discovering decent theories. Concomitantly, one is not free to impose arbitrary additional metaconstraints (e.g. computational plausibility) which could conflict with the well-defined basic goals of the enterprise. (Prince and Smolensky, 2004 [1993]: 215f., emphasis in original)

Since its inception, generative grammar indeed followed this line of thought: Chomsky has always argued that competence is not about implementation, and that implementational arguments have no bearing on the properties of the model of competence that linguists are supposed to develop. The minimalist programme (Chomsky, 1995, 2000: passim), however, clearly suspends this perspective: grammar must respond to implementational requirements. The whole point of the minimalist approach is to make grammar evolve in response to extra-grammatical factors, that is, interface requirements. Phase Theory, for example, cuts the computation of a full sentence into independent pieces for reasons of computational economy regarding the limited availability of active memory, a costly cognitive resource (e.g. Chomsky, 2000: 101; 2001: 15).

In this environment, OT cannot hide behind competence anymore in order to escape the discussion of computational complexity. The greater generality of constraints with respect to rules and their dissociation from a specifically phonological vocabulary is thus a problem. On the other hand, it has crucially contributed to the success of OT and parallel computation. De Lacy (2007b: 14ff.), for example, recapitulates the advantages of parallel computation over serialism: ordering paradoxes, global conditions, conspiracy. The question, then, is whether the generalizing remedy (candidate computation by prose-based constraints) is not worse than the original disease (time-consuming serial application of rules).

5.4 Tabula Rasa in 1993 and Hybrid Rule-Constraint Computation?

A worrisome property of the OT literature and of the self-understanding of OT is the constant attempt to assess a tabula-rasa interpretation of the year 1993: de Lacy (2007b), for example, is much concerned with showing that OT is different from anything that phonology has produced before 1993 in general, and with making OT antithetical with respect to SPE in particular (de Lacy, 2007b: 13, opposes OT to ‘the dominant theories before OT – SPE and its successors’). The tabula-rasa claim may help to assess a new theory, to raise funds and to gain academic positions, as it did in the 1960s when generative phonology rolled over the structuralist establishment – but it is as wrong now as much as

it was then (see Goldsmith, 2008). Even a cursory look through Anderson's structure-process prism is enough to be convinced: OT takes phonology way back to the 1960s when representations did not exist and computation was king (more on this in Section 5.6). Van der Hulst and Ritter (2000) describe the SPE-heritage of OT in greater detail. Also, it was shown in Sections 5.1 and 5.2 that OT roots in autosegmentalism-promoted well-formedness conditions. Also, Prince and Smolensky's (2004 [1993]) original manuscript does not really license the picture that present-day OT-based story telling tries to establish.

This is not to say, of course, that there is nothing new in OT: our conception of computation has profoundly changed since OT and other theories laid the focus on constraints. The merit of OT is the application of parallel computation to phonology, as well as the promotion of the specific view whereby constraints are ranked and violable.

In this context, it is interesting to compare the evolution of computation in syntax and phonology. As Hale and Reiss (2008: 202ff.) point out, the Pisa turn syntax and the movement against extrinsically ordered rules in phonology were not only abolishing the same kind of serialism in both areas, and were not only doing away with rules. They also replaced rule-based serialism by *free generation-cum-filters*. In GB syntax, morphemes could 'freely' be concatenated, and Move- α could 'freely' apply. Ill-formed results were then filtered out by global constraints (on locality, case etc.). The minimalist version of this conception of computation focuses on the interfaces, PF and LF, which impose conditions that make the derivation either crash or converge. Constraint-based computation in phonology follows the same track: in OT GEN does the free generation, and constraints filter out the optimal candidate. Unlike in syntax where filters have equal rights and are inviolable, however, constraints are ranked and violable in OT.

Given this overall picture, an interesting question is certainly whether the correct solution for computation is only binary, that is, either rule-only or constraint-only. Calabrese (2005), for example, holds the principled position that a sound theory of phonology must have both serially ordered rules and constraints: while the former are instructions to create a given configuration, the latter specify which configurations must be avoided (constraints apply only to Kisseberth's 1970a conspiracies). This is indeed the naturally grown state of the art of the 1980s (with enlarged competences for well-formedness constraints, which may also concern purely melodic configurations): recall from Sections 4.2 and 5.1 that well-formedness constraints (i.e. output filters) are a direct consequence of (autosegmental) representations and peacefully cohabitated with traditional rule-based computation. Hale and Reiss (2008: 209ff.) offer instructive discussion of the rules-cum-constraints option (concluding that both hybrid and constraint-only systems are inaccurate: computation must be purely rule-based in their opinion).

5.5 OT and its Connectionist Endowment

It was briefly mentioned in Section 2.1 that generative grammar is the linguistic outgrowth of the cognitive revolution of the 1950s, which was anti-behaviourist, anti-empiricist and dualistic in nature. Connectionism is the modern version of the reverse intellectual position. Prince and Smolensky (2004 [1993]: 217ff.) provide an informative discussion of the relationship of OT with connectionism. They argue for cherry-picking: while parallel computation is taken over, other major connectionist tenets are rejected. For example, Prince and Smolensky do not accept the connectionist anti-symbolic stance that rejects symbolic representations (such as syntactic trees) because the only relevant level where decisions are made is neuronal: in a connectionist environment, neurons only react on activation levels, hence cannot parse, or distinguish, between symbolic objects (e.g. Fodor and Pylyshyn, 1988; Dinsmore, 1992).

Just like the standard theory of Cognitive Science, OT recognizes a symbolic level of representation. The place for connectionist non-symbolic ('colourless') computation, then, is an intermediate level between the symbolic level and the physiologically neural functioning of the brain. This conciliatory position that rejects reductionism (i.e. the denial of the mind as an independent level of analysis) and where the connectionist level mediates between the mind and the brain is defended by Paul Smolensky since his earliest work (Smolensky, 1986, 1987, 1988, 1991) and up to the present day (Smolensky and Legendre, 2006).

Also, Prince and Smolensky reject the neo-behaviourist take of connectionism regarding acquisition according to which 'knowledge of language can be empirically acquired through statistical induction from training data' (Prince and Smolensky, 2004 [1993]: 217).

Although Prince and Smolensky (2004 [1993]) do not mention this issue, OT obviously also rejects the connectionist claim that there is no difference between computation and storage. In the connectionist perspective, the 'experience' of a neural network – the equivalent notion of memory – is acquired when the patterns of connectivity change: neurons may develop new connections (synapses), may lose old connections or modify the strength (weight) of existing connections (the two former are often viewed as a special case of the latter). The computational units themselves have no variable behaviour that contributes to the properties of the whole, which is exclusively determined by the connective network (see, for example, Stillings et al., 1995: 114ff. on connectionist models of memory).

All linguistic theories since Antiquity of course rely on the assumption that there is a lexicon that exists independently of grammatical activity which transforms lexically stored objects into actual speech. The linguistic mirror of the connectionist non-separation of storage and computation is so-called 'Cognitive' Grammar,²⁴ which was founded by Langacker (1987) and is overtly

empiricist, (neo-)behaviourist and anti-generative (see Taylor, 2002): Langacker (1987, Vol.1: 42) talks about the 'rule/list fallacy'. The phonological offspring of this line of thought is represented by exemplar- and usage-based approaches in general, and by Joan Bybee in particular, who writes that

[I]n linguistic regularities are not expressed as cognitive entities or operations that are independent of the forms to which they apply, but rather as schemas or organisational patterns that emerge from the way that forms are associated with one another in a vast complex network of phonological, semantic, and sequential relations. (Bybee, 2001: 20f.)

Finally, another important connectionist headline is the aforementioned PDP, which contrasts with the regular assumption that computation is serial. While on the count of standard Cognitive Science the output of one computation is the input to another, multiple computations take place in parallel in the connectionist perspective. Also, the units that carry out computation – neurons, or clusters thereof – are not specialized for any particular computational task, or for a particular input material (computation is non-symbolic and not specific to any domain). Rather, neurons are all-purpose computational units that are able to perform any computation on the grounds of any type of information submitted. This is why connectionist computation is called *distributed*.

A corollary of distributed computation is the claim that there are no specialized computational units made of clusters of neurons that can be delineated in the brain: computation is opportunistic and does not need any specialization of its support units, the neurons. The modular theory of the mind (Fodor, 1983), which incarnates as the inverted T model in the generative architecture of grammar, is based on the reverse assumption: there are stable, genetically endowed, content-sensitive (i.e. symbolic) computational units that are devised to a very narrow and specific function, which can only work with a particular type of input vocabulary (domain-specificity), and can do nothing else than what they have been designed for.

In sum, the connectionist perspective may be characterized by the fact that computation is content-free: all other properties follow from this assumption. That is, the mind does not know what it is doing when computation takes place: computation is only general-purpose, that is non-specialized for any task or function; it works without reference to any symbolic code, which would make the operations specific to a particular domain or content since symbols are symbols of something, and may be opposed to symbols of a different kind.

The question is whether the cherry-picking in the densely interrelated network of connectionist assumptions that Prince and Smolensky (2004 [1993]) propose is viable: the genetic code of OT rejects basically all tenets of connectionism save one, parallel computation. Parallel computation is represented by the two Ps in

PDP (Parallel Distributed Processing), but Prince and Smolensky do not address the question of the D, which is anti-modular. We will see in Section 5.7 that the D appears to be a direct consequence of parallel computation: it is constantly working on OT practice (if without explicit discussion) and has induced what I call the scrambling trope, that is, the creeping dissolution of modular contours. The same holds true for content-free computation, which has made representations irrelevant and interchangeable, before dissolving them in computation altogether (representations are 'emergent', rather than given). This computation-trope is discussed in the following section.

The conclusion, then, is that parallel computation has probably entered the generative paradigm with some more empiricist luggage, and the question for further development is a theory that maintains the rationalist and anti-empiricist core of generative grammar while also implementing constraint-based and parallel computation.

5.6 Grammar Reduces to Computation: Representations Are Demoted to Decoration

Knowing about the connectionist roots of OT helps to understand the extreme computational orientation that phonology has taken under its lead since 1993. It is sometimes rightly recalled that OT is a theory of constraint interaction, not of constraints. This means that OT does not supply any substance itself: there are genuine vocabulary items in structuralism (phonemes), SPE (segments) and autosegmental theory (autosegmental structure), but there are no OT-specific representational items. OT uses whatever representational material is available, and may well produce the same result with entirely different (and incompatible) vocabulary.²⁵ It is difficult not to establish a direct relationship between the fact that OT is a purely computational theory where representations make no sovereign contribution to the definition of grammaticality (which is decided by constraint interaction alone) and its content-free connectionist prototype.

Paul de Lacy's (2007a) edited handbook, and de Lacy's (2007b) introduction to the volume in particular, document the global trend from representation to computation in much detail. It is hard to find a thematic chapter of the book that does not insist on this evolution. A good example is John Harris's 2007 contribution on representations. The author, who is not exactly known for anti-representational positions, defends the categorical, that is, non-gradient character of representations. In the end, however, the reader learns why it is that representations can be non-gradient at all: because they do not exist. The world, including grammar, is definitely gradient and computational: categorical objects can exist at best as a product of gradient computation (structure is emergent, rather than given: 'categorical behaviour is emergent rather than an

inherent property of these descriptors', Harris' conclusion, p. 137). Rather than write about representations, Harris thus ends up talking about 'descriptors' since this is the only role that they can play in a world where decisions are only made by constraint interaction: representations 'still have an important heuristic value as descriptors to be used in the building and experimental testing of models of phonological grammar' (Harris, 2007: 137). In other words, representations are decoration: they may help the linguist to picture the result of constraint interaction – but they have no impact on grammar at all.

This is indeed the conclusion that follows from the OT-tenet that the *only* means to determine grammaticality is constraint interaction. Hence, whatever items of the representational furniture of the 1980s are used, they are mere decoration that do not function as sovereign arbitrator, and do not have anything to say regarding grammaticality (a structure with line-crossing, for example, may be the optimal candidate if all other candidates violate higher-ranked constraints).

Given Stephen Anderson's prism and the constant see-saw movement observed, a fair question to be asked in the face of the new OT-loop back into the 1960s is *why* theories go down the representational or the computational road. When autosegmental representations were developed, the motivation was clear: gain of insight (tone spreading, the possibility to characterize the coda disjunction __{#,C} as a single phonological object etc.) and the promise of an efficient instrument against the plague of overgeneration. De Lacy (2007b) examines the question why OT has progressively replaced representations by computation. The answer is more or less that representations are eliminated because their function *can* be taken over by computation: 'OT has allowed the burden of explanation to move from being almost exclusively representation-based to being substantially constraint-based' (24). The question why there *should* be such a movement (which, recall, is not at all what Prince and Smolensky, 2004 [1993] had originally envisioned) is left unanswered: we do X not so much because we want to and have good reason to do it, but simply because we *can* do it. In other words, OT is a theory of computation and promotes what it is competent for without looking left or right. The progressive elimination of representations is thus but a side effect of this computational trope, which roots in the decision that the *only* thing that determines grammaticality is constraint interaction.

This proviso, however, does not follow from OT: OT is a theory of parallel computation that uses ranked and violable constraints. This does not lay any claim on how much of the explanative pie is computational: the view that 100% has been installed without discussion or comment and today is part and parcel of OT. Yet it is only one possible attitude. Another view is expressed in a small but growing body of literature to which Marc van Oostendorp (van Oostendorp, 2002, 2003, 2005, 2006) has contributed a good deal, and which is condensed in

Blaho et al., (eds.) (2007); Optimal Domains Theory belongs to the same family (Cole and Kisseberth, 1994; Cassimjee and Kisseberth, 1998). Blaho et al., (eds.) challenges Freedom of Analysis: you ought not to be free to do what you want with representations. In terms of the classical OT grammar, this means that there are restrictions on GEN, which produces only a subset of logically possible candidates.

The idea that OT is a complete theory of grammar has been tacitly entertained since its inception. The bare existence of versions of OT that place restrictions on GEN show that this view is overstated: OT is not a theory of grammar; it is a theory of a piece of grammar, computation. Anderson (1985) says that it takes more than just computation to make a grammar. The least common denominator of OT, then, is parallel computation that uses ranked and violable constraints. All the rest is free and a matter of choice of the analyst, who may or may not be a generativist, may or may not be a functionalist, may or may not assume a modular architecture, may or may not be representationally oriented, may or may not believe in the virtue of serial ordering of phonological (and/or grammatical) events (see below), may use this or that representational system, and so forth (this issue is further discussed in Scheer, 2009).

5.7 The Scrambling Trope

Closely related to the trope for computation is the pervasive reflex of OT to make distinct things indistinct; that is, to put them in the same constraint hierarchy, to intersperse them and to assess them in one go. Of course there are many different degrees of scrambling, but the tendency is clearly towards a single grammatical space where anti-derivationalism is enforced globally. The most visible result of this trope is anti-cyclicity, that is, the rejection of inside-out interpretation because of its serial character (e.g. Kager, 1999: 277). OT has produced a whole anti-cyclicity literature (which has an important intersection with the anti-opacity literature) that proposes alternative, strictly parallel ways of communicating with morpho-syntax: co-phonologies, indexed constraints, OO (Output-Output) faithfulness and so-called interface constraints.

Three more diagnostics for OT's misty (and largely unreflected) relationship with modularity may be mentioned. First, uncontradicted violations of Indirect Reference (i.e. the prohibition to make reference to untranslated morpho-syntactic categories that was established by Prosodic Phonology) are commonplace in OT. That is, ALIGN and WRAP constraints make constant reference to morpho-syntactic structure and labels; interface constraints such as FAITH-root and FAITH-affix make reference to designated morpho-syntactic categories (even reference to individual morphemes is not a problem, Anttila, 2002, provides an overview) and thereby reincarnate the SPE-practice of supplementing

rules with morphological diacritics. Second, mapping (of morpho-syntactic into phonological prosodic categories) is done *in* the phonological constraint hierarchy (rather than outside of the phonology as was the case in Prosodic Phonology): ALIGN and WRAP are interspersed with purely phonological constraints. Finally, constraints whose formulation combines phonological and morphological instructions are customary (see Yip, 1998, on this issue).

Kingston (2007) discusses the scrambling trope on the example of the abandon of the distinction between phonetic and phonological constraints. He also points out the causal relationship between the move from serial to parallel computation on the one hand and the everything-is-the-same programme on the other:

[r]eplacing serial derivation by parallel evaluation removes the barrier to phonetic constraints being interspersed among and interacting with phonological constraints. [. . .] Future research will determine whether phonological and phonetic constraint evaluation are a single, integrated process, as advocated by Steriade and Flemming or instead sequential, as advocated by Zsiga. (Kingston, 2007: 432)

There is thus reason to believe that the commitment to non-serialism is the driving force behind OT's scrambling trope. Despite this in-built tendency, however, adhering to modular-destructive indistinction in grammar is a personal choice of the analyst, not a fatality. Those who argue with anti-derivationalism in order to challenge cyclic derivation and to set up a single constraint hierarchy where phonetic, phonological, morphological and even syntactic constraints (sic, Russell, 1999) are interleaved make a category mistake: derivation and computation is not the same thing (see Scheer, forthcoming a). OT is committed to parallel *computation*, and in generative grammar the unit where computation takes place is the module. Grammar is made of several modules, each with a distinct computation that works on distinct (i.e. domain specific) vocabulary. Hence, nothing withstands a perspective where all linguistic computation is perfectly parallel, but distributed over distinct and serially ordered computational systems (modules). It is only when the non-derivational claim is laid on the entire grammar that the scrambling trope appears.

That one can resist the scrambling trope is shown by the fact that there are derivational versions of OT that allow for serial communication among modules while observing strictly parallel modular-internal computation: Derivational Optimality Theory (DOT) (Rubach, 1997: passim) and Stratal OT (Kiparsky, 2000; Bermúdez-Otero, forthcoming a) are reincarnations of Lexical phonology in the new constraint-based environment.

Beyond the diluting of modular contours in one single constraint chamber, another target of the scrambling trope is the lexicon. Richness of the Base prohibits the introduction of any distinction in the lexicon that goes beyond melodic contrast (this is one of a number of readings, the literature offers various interpretations) because GEN must not be marshalled in any way when creating candidate variation. Hence, nothing beyond segmental contrast can be hard-wired in the lexicon. This is why input forms in OT typically look like underlying forms in SPE: they are made of a linear chain of segments. In practice, this means that the progressive integration of constituent structure into the lexicon that was undertaken in the 1980s (see Section 4.4) is ruled out, and the corresponding labour is transferred to the constraint chamber. But even more generally there is a trend in OT to introduce more and more unpredictable information by constraint. This movement is documented, for example, by de Lacy (2007b: 19f.); it contributes to the computationalization of grammar, and is also a step in the direction of the connectionist indistinction between computation and storage (whose linguistic expression, recall, is Langacker's 1987 Vol.1: 42 and Bybee's 2001: 20f. 'rule/list fallacy').

Finally, two other types of static information are also expressed by constraints, rather than by some hard-wired (lexical) recording: parameters and inventories. The transformation of the lexical recording of language-specific parameters into a variable constraint ranking (children do not set parameters on-off, but rank relevant constraints) is one of the most basic ambitions of OT, and also one that is generally held to be an area where OT is successful: cross-linguistic variation is described in terms of different rankings of the same set of universal constraints (factorial typology).

Regarding inventories, the idea is to kill two birds with one stone: inventory properties and some of the language-specific processing are derived from the same source, that is, a set of constraints. Steriade (2007) provides a well-informed overview of this project, which is also developed in *Government Phonology* since the mid-1990s (where the instrument are so-called Licensing Constraints, see Charette and Göksel, 1994, 1996; Kaye, 2001). In my opinion, the computationalization of parameters and of inventory information has a different quality when compared to other aspects of the trend towards scrambling: it does not destroy any distinction that is central to generative thinking (as is the case with modular contours), nor does it undo analytic advances (as is the case with the prohibition of constituent structure in the lexicon); rather, it gives a different expression of something that must be stated somehow anyway (parameters), or provides a new analytic perspective for the (belated) reconciliation with the heart of structuralist thinking (inventories and systemic pressure, which may also be at the origin of phonological processes).

In conclusion, it appears that all letters in PDP (Parallel Distributed Processing) make sense and are interrelated. This casts doubt on the idea that

individual connectionist tenets may be cherry-picked: recall that Prince and Smolensky (2004 [1993]) import parallel computation into OT, but do not mention its distributed aspect. It looks like the distributed character of computation, that is, the connectionist everything-is-the-same programme, is a direct (perhaps even ineluctable) consequence of parallel computation.

6. Counterrevolution #2: Neo-Serialism and Representations

6.1 Serialism *within* Phonological Computation: OT-CC

The latest evolution within OT has initiated the second counterrevolution: John McCarthy was at the forefront of the anti-derivational movement in the 1990s and 2000s. In his 2007 book (McCarthy, 2007), he makes a radical about-turn and now promotes serialism. His OT-CC (Candidate Chains) has an entirely different quality than OT versions of Lexical Phonology (DOT, Stratal OT) where the relationship between modules is serial, but phonological computation itself strictly parallel: OT-CC holds that phonological computation itself is done on the basis of serially ordered instructions. This does not mean, however, that parallel computation is abandoned, or that McCarthy goes back to ordered rules. Rather, OT-CC is a version of Harmonic Serialism (see Section 5.2, and actually is more and more referred to under this name): the overall computation of an input is cut into a series of different evaluations where the output of EVAL is fed back into GEN, which adds harmonically improved candidates on each round. Candidate chains thus reflect the chronology of computational events (like intermediate forms in a serial SPE-derivation). EVAL is strictly parallel and invariable; it evaluates candidate chains, rather than single candidates, and so-called PREC constraints, a new constraint family which is the actual locus of ordered instructions, assign violation marks on the grounds of the comparison of different members of the candidate chain (PREC constraints react on the order of faithfulness violations).

There is an obvious parallel to computation as conceived of in Government Phonology (see Section 5.2) where input strings are also modified by a fixed set of instructions (which however are unordered and unranked) in iterative fashion: the output of computation also loops back in order to become the new input, and the derivation also ends when the string cannot be further modified ('harmonically improved').

OT-CC is the capitulation of anti-serialism in the face of opacity: John McCarthy has tried hard to find a solution for opacity within parallelism (Sympathy Theory, Comparative Markedness), but was not any more successful than the half dozen of other opacity killers that can be found in the literature (e.g. targeted constraints, enriched inputs, F&F conjunction, M&F conjunction,

turbidity). McCarthy has studied one particular case of opacity (Bedouin Arabic) in depth to make sure that there is no lexical, morphological or other parasitic conditioning, and also that the processes at hand are really synchronically active and productive. He then argues along the lines of the strongest case: one such language is enough to force phonological theory into serialism.

During the anti-derivational period of the 1990s and 2000s, extrinsically ordered rules were upheld by a body of work including Vaux (2003), Calabrese (2005), Halle and Matushansky (2006), Raimy and Cairns (eds., forthcoming) and Hale and Reiss (2008). These thus share with OT-CC the original tenet of generative grammar that (phonological) computation is serial and requires an extrinsically ordered set of instructions. The debate, then, is not whether we need serial computation, but rather how serialism should be implemented: as a looping and hybrid mechanism that also relies on parallel computation (OT-CC), or as the classical system that uses extrinsically ordered rules but does not provide for loops or parallel elements.

6.2 Representations: You Cannot Compute Nothing

The other strand of the counterrevolution are representations that are not just mere decoration. Meaningful representations contribute a sovereign and unoutrankable arbitration role. That is, their verdict is absolute and not subject to any further evaluation: an ill-formed representation is ill-formed no matter what and independently of constraint interaction. An ill-formed structure where, say, association lines cross is out; it cannot turn out to be the optimal candidate because all other candidates violate some higher-ranked constraint. An ill-formed representation is either repaired, or the derivation crashes. In other words, representations that deserve their name break with the OT-baseline according to which the only way to determine (relative) grammaticality is constraint interaction.

Another property of meaningful representations is that they are not necessarily the result of computation: structure may be given, and neither its genesis nor its role as arbitrator has to be 'emergent'. Representations of this kind are also meaningful because they are not interchangeable. It was mentioned in Section 5.6 that OT is a theory of computation and therefore has no opinion on representations: it has not developed any and can (actually does) work with any furniture of the 1980s, including competing and incompatible structure (see note 25). Constraint interaction will always squeeze out the right result. The effect is that representations first become decoration, and then completely disappear.²⁶ Since representations are irrelevant, they have stopped being developed: there is hardly any discussion regarding the comparative merits of

competing representational theories ('I do this analysis using morae, but if you prefer x-slots that will do as well, you will just have to change the constraints accordingly').

Theories that develop meaningful representations include Government Phonology (Kaye et al., 1990; Charette, 1991; Harris, 1994, 1997; Kaye, 2005; Pöchtrager, 2006), Dependency Phonology (van der Hulst and Ritter, 1999; Botma, 2004; van der Hulst, 2005), Substance-Free Phonology (Hale and Reiss, 2008; Blaho, 2008) and the aforementioned rule-based approaches.

Government Phonology is especially often regarded as a representation-oriented theory. It has continuously worked on the development of representations since the mid-1980s, namely in the area of syllable structure. The genuine idea defended is that constituent structure is lateral, rather than arboreal: in Standard Government Phonology (see the references in the preceding paragraph), whether or not a consonant can be a coda depends on the availability of licensing from a following onset. That is, *r* in *Vr.tV* can be a coda because a licensing onset to its right is available, but word-final consonants cannot be codas since there is no following onset: therefore *r* in *Vr#* is an onset (followed by an empty nucleus).

So-called CVCV (or strict CV, Lowenstamm, 1996; Scheer, 2004; Szigetvári, 2002; Szigetvári and Scheer, 2005; Cyran, 2003), a development of Standard Government Phonology, takes the lateral idea to its logical end: instead of a hybrid system where arboreal and lateral structure cohabitate, constituent structure boils down to a strict sequence of non-branching onsets and non-branching nuclei (no branching constituents, no rhymes, no codas). In this environment, the coda behaviour of a consonant is a consequence of the fact that it is followed by an empty nucleus which is unable to license its onset because it is empty (consonants that show onset behaviour are followed by filled nuclei that is hence able to dispense licensing).

The flat, that is, non-arboreal constituent structure that is the result of strict CV offers an in-built explanation why there is no recursion in phonology: the absence of arboreal structure witnesses the fact that phonology has no access to Merge (or an equivalent tree-building device); in absence of Merge, however, there can be no recursion (Scheer, 2004: xlix ff., Neeleman and van de Koot, 2006).

Finally, it was already mentioned that there is also a representation-rehabilitating strand within OT that was condensed in Blaho et al., (2007) (see also Optimal Domains Theory, Cassimjee and Kisseberth, 1998). Van Oostendorp and van der Weijer, (2005), for example, make the argument based on vocabulary (recall domain-specificity that is required by the modular architecture): OT needs a universe of discourse. That is, you cannot compute nothing (or interchangeable phantoms). Regarding melodic representation, Hall (2001) and Clements (2001) may certainly be counted as expressions of this line of thought as well.

7. Conclusion

I try to condense in (2) below what I take to be the important issues and the true advances that have been made in the field and that will probably still play a role in 50 years.

(2) *Important issues*

- a. the system
systemic pressure on phonological events and their implementation into a formal system
- b. melodic primes
binary, privative or a blend thereof?
- c. type of relatedness of alternating forms
 1. computational: by a phonological process (common underlying form) or by a non-phonological process (allomorphy, suppletion, analogy)
 2. non-computational: two separate lexical entries
- d. phonological computation
 1. type of computational instruction
rules (specifying the modification of the input and its triggering conditions) or constraints (general requirement or prohibition with no modificational instruction)?
 2. serial, parallel or a blend thereof? If serial, what kind of serialism: extrinsically ordered rules, harmonic serialism? If parallel, what kind of parallelism: hard or violable constraints, if violable, dominance expressed by weight or by rank?

None of these issues is settled, and I doubt that (2c) will ever be. For (2a) is not about whether or not phonological processes are under the spell of systemic pressure: hardly anybody doubts that they are; the question is how exactly this could be expressed in a formal system.

The list under (3) below gives what seem to me are true achievements and advances of generative phonology.

(3) *Achievements and advances*

- a. cyclic derivation
- b. modular architecture of grammar
- c. interactionism

- d. autosegmental representations, that is, which can be ill-formed, act as arbitrator and are not necessarily the result of computation
- e. privative melodic primes
- f. constraints and parallel computation.

This list may be divided into two groups. On the one hand, cyclic derivation and the modular architecture of grammar are the generative baseline: they were present upon inception of the generative enterprise, and they concern architectural properties of the grammar that lie beyond narrowly phonological concerns. Also, while modularity is shared with structuralism (Level Independence), cyclic derivation is a genuinely generative discovery. It is difficult to see how anything could be called generative that does not endorse these two properties. It was shown that OT in its present state freely violates modularity, and that this appears to be the application of the D (distributed: the scrambling trope) of PDP, which itself is a direct consequence of the P (parallel computation). It remains to be seen whether Prince and Smolensky's (2004 [1993]) cherry-picking (we are true generativists and hence reject the empiricist freight of connectionism, except for parallel computation) can bear fruit.

The other group of the list under (3) encompasses items that have appeared in the course of the development of generative theory. Two of them, (3e–f) also appear in the list of issues: there is still debate whether they exist at all and some believe they do not, but in any case the existence of this debate alone and the issues raised are valuable progress. While constraints are used by all theories in one form or another (recall that well-formedness conditions of the 1980s such as the OCP are a form of constraints), it is true that strictly rule-based approaches may reject the idea of parallel computation altogether. In any event, the parallel idea is an entirely new perspective on computation that has injected fresh air into the settled waters of the serial model of Standard Cognitive Science in general, and of its linguistic outgrowth in particular. Note that unlike modularity, serialism is not an essential of the rational conception of the mind: it is simply the default of the 1950s that was introduced by the Turing/von Neumann machine. All other things being equal (and this is the question raised by cherry-picking), computation could be parallel as well. Generative syntax has shown that non-serialism is possible within the generative paradigm (recall, however, that the constraints on syntactic well-formedness are hard, rather than ranked and violable).

Finally, the two other items of the second group, (3c–d) are the major scientific advances of generative phonology in my opinion. Again, one concerns an architectural property: interactionism was invented by phonologists in Lexical Phonology but has made a brilliant career in syntax: it is the spine of current minimalist thinking (under the name of derivation by phase). The other item is

the face of modern phonology: autosegmental representations are certainly the least common denominator of all currently practiced theories, even if they are demoted to pure decoration in (some versions of) OT.

This record appears to be quite positive and encouraging, an impression that needs to be tuned down, though, when we consider that generative phonology seems to go in circles with respect to some fundamental issues, and spends quite some time reinventing the wheel. The systemic question was thrown over board in the 1960s without argument, and reappeared on the research agenda only recently. But worse is the situation regarding Anderson's (1985) concern for a reasonable balance between representation and computation. Generative phonology has made a complete loop (which is also described by van der Hulst 2004: computational SPE → super-representational 1980s → super-computational OT) and is maybe engaging into another round (cf. the still timid return of representations within and outside of OT). The parallel with syntax in this area is quite striking: like phonology, syntax was anti-lexicalist and very computationally oriented in the 1960s, explored the lexicalist and representational track in the 1970s and 1980s, but has gone back to extreme anti-lexicalism and proceduralization since the turn of the minimalist programme.

It would be nice if phonology (and syntax for that matter), like adult science, could be said to follow a linear trajectory from less to more knowledge, in a cumulative movement that builds on and learns from the experience and errors of the past. This has certainly not been the case thus far (regular assertions of the contrary from OT quarters notwithstanding), but maybe the see-saw movement can be broken this time: we do not know where exactly the red line runs, but there is good reason to believe that both representations and computation are needed in order to make a grammar.

8. Notes

1. See, for example, Newmeyer (1986, 1996), Anderson (1985), Encrevé (1997), Goldsmith (2005), Goldsmith and Laks (2000, forthcoming), Koerner (2002a, 2004), van der Hulst (2004), Laks (2001) and the five contributions to Auroux et al. (2006) regarding generative phonology by J. Durand, B. Laks, J. Goldsmith and M. Noske, J. Jensen and J. Lowenstamm.
2. The revolutionary metaphor is commonplace in linguistic historiography and, as Koerner (2002b) points out, part of generative story-telling. Koerner dwells on the (in)adequacy of this reference at greater length, also in regard to Kuhn's model.
3. A fair debate is about whether we are talking about sub-modules, or just about an aggregation of independent modules whose collaboration happens to produce language. The same issue arises in other areas of the cognitive system: we know today that vision, for example, is made of different sub-systems that compute colour, shape, face recognition and movement independently.

Another aspect of the debate regarding language-internal computational units is what Chomsky (1981) has called the sub-systems of GB modules (case theory, theta

- theory etc.). Space restrictions preclude further discussion (but see Scheer, forthcoming a).
4. Syntacticians usually do not go beyond statements such as ‘and then PF takes over’. In the recent minimalist environment, they actually use PF as a dustbin for things that they want to get rid of in order to have a ‘clean’ syntax (e.g. deletion of words or even entire phrases ‘at PF’). The resulting ‘dirty’ phonology, then, is not anything that they are much concerned with (see Scheer forthcoming a).
 5. Whether syntax and morphology are the same or two distinct computational systems is a debated issue that is orthogonal to the present discussion.
 6. Bermúdez-Otero (forthcoming b) provides an informed overview of cyclicity and the role it played in the development of generative phonology (see also Scheer forthcoming a).
 7. Though not for all: today substance-free phonology (Hale and Reiss, 2008; Blaho, 2008) holds that substance, that is, melodic properties of segments, is irrelevant for phonological computation: substance is the subject matter of phonetics (if phonology talks about substance, it is liable for substance abuse). Hence regarding melody, anything and its reverse can be turned into anything and its reverse in any context and its reverse. So-called crazy rules (Bach and Harms, 1972) are a relevant factor in the discussion (Scheer, forthcoming b).
 8. This principle has later been made explicit in Government Phonology: ‘non-arbitrariness: There is a direct relation between a phonological process and the context in which it occurs’ (Kaye et al., 1990: 194).
 9. The process whereby *k* is turned into *s* (before *-ity*) is called velar softening. It is emblematic of the discussion of derivational relatedness and has therefore aroused a substantial body of literature that includes Chomsky and Halle (1968: 219ff., 426ff.), Kiparsky (1982a: 40f.), Kaye (1995: 312, 328), Halle (2005) and McMahon (2007). Hayes (1995b) and Green (2007: 172ff.) provide a documented overview. On the structuralist side, Harris (1951: 221ff.) argues in favour of allomorphy.
 10. Vennemann (1976 [1971]) is quoted as a 1971 manuscript by Hooper (1975) and Zwicky (1974). Even though Stampe’s Ph.D is typically referred to as an important source in the NGP literature, it appears that the development of NP and NGP was parallel, rather than based on a common ancestor.
 11. See Newmeyer (1998) for an overview of what he takes to be the two basic approaches to linguistics, formalist and functionalist thinking. There is no place to further discuss the differences between the two Natural Phonologies. Laks (2006) offers more material on this issue (see also Hooper, 1975: 544ff.) and discusses their further evolution: while NGP has no modern offspring (Joan Hooper, who today publishes as Joan Bybee, is engaged in a usage-based, that is, empiricist approach, and thus has left generative grounds altogether), NP continues to be actively developed (see the references quoted).
 12. Cole (1995) and Bermúdez-Otero (forthcoming a) provide a good overview of this movement, which is also explained by Kiparsky himself in Kiparsky (1982b, 1993), and by Anderson (1981: 530ff.).
 13. Today Trisyllabic Shortening is typically not considered a synchronically active process: it faces quite a number of counterexamples such as *obese* [ɔwbiis] – *obese-ness* [ɔwbiisnəs], *obese-ity* [ɔwbiisit]. Hayes (1995) and Green (2007: 172ff.) provide an informed review of its status.
 14. Cole (1995: 72) discusses the two distinct origins of Kiparsky’s SCC at greater length, and Scheer (forthcoming a, in press a) follows the development of all no look-back devices from Chomsky (1973) until the modern syntactic Phase Impenetrability Condition. Finally, it is to be noted that the marriage of Chomsky’s SCC with derived environment effects was first proposed by Halle (1978).

15. On the syntactic side, the precursor of the PIC was Riemsdijk's (1978: 160) Head Constraint, which does the same labour as the PIC, but not for the same reasons (see Scheer, forthcoming a).
16. Giegerich (1999: 7ff.), McMahon (2000: 35), Rubach and Booij (2003) and Bermúdez-Otero (forthcoming a) offer well-informed historical overviews (see also Scheer forthcoming a).
17. This is only a rough and incomplete formulation of the phonological ECP, which also does not mention the central notion of government; see Gussmann and Kaye (1993), Scheer (2004) for details.
18. The linear solution was to add extra items to the vowel inventory, *ĩ* and *ĩ̃*, that were then either transformed into [ɛ] ('lowered') or deleted, that is, subject to absolute neutralization.
19. Constraints as such, however, are probably as old as linguistic analysis: so-called morpheme structure constraints defined what a possible morpheme is in SPE, and structuralist as much as neogrammarian thinking appealed to restrictions, requirements and prohibitions. Van der Hulst (2004) provides an overview of the generative record regarding constraints.
20. There was an attempt to do away with extrinsically ordered rules in the 1970s around the work of Andreas Koutsoudas (Koutsoudas et al., 1974; Koutsoudas, 1976a): Koutsoudas claimed that there is no empirical basis for extrinsic rule ordering; its effects can be derived (and predicted) from other factors. The evidence presented was not judged convincing, though, and the proposal did not have any posterity. In any event, it was entirely unrelated to the representationally dominated environment of the 1980s in general and the 1980s antipathy against derivationalism in particular.
21. Bromberger and Halle argue along the lines of the strongest case, which they claim is the comparison of the well-attested phenomenon of Canadian Raising in Canadian English (dialect A) with another Canadian dialect (dialect B) whereby the only difference is the order of application of two rules. Unfortunately, the only source for dialect B is a three-page article by Joos (1942), whose informants were among his pupils in an Ontario public school. Despite extensive study of Canadian varieties by Canadian dialectologists such as Chambers (1973), though, no trace of dialect B could be found some 30 years later. Despite the arguments of Kaye (1990c, 2008), the Canadian raising case continues to spook though the literature as alleged support for rule ordering.
22. Mohanan (2000: 146) argues that they do not because one can always be translated into the other. The point, however, is not whether the 'correct result' can be achieved; rather, *the way* in which this result is achieved matters, and is different. See the informative discussion in Hale and Reiss (2008: 195ff.) on this issue.
23. The only exception that I am aware of is van Oostendorp and van de Weijer's (2005) attempt to define what they call a universe of discourse for the expression of instructions in OT.
24. I use quotation marks in order to refer to this framework because its name fallaciously suggests that it has a copyright on cognitive aspects of grammar, and that anything which is non-Langackerian must be non-cognitive.
25. For example, Lombardi (2001a: 3) writes with respect to melodic representation: 'the tenets of OT, regarding constraint violability and ranking, make no particular claims about phonological representations. We could, for example, do OT with any kind of feature theory: SPE feature bundles or feature geometric representations, privative or binary features, and so on.'
26. An eloquent example is the 'representation' of melody that is currently practiced in OT: binary features are about as sophisticated as they were in SPE (i.e. an amorphous, unarticulated set), and constraints that manipulate melody boil down to 'star [feature]' (e.g. *dorsal etc.). Clements (2001) and Hall (2001) bemoan this evolution.

Bibliography

- Adger, D. (2007) Stress and phasal syntax. *Linguistic Analysis* 33: 238–266.
- Aichert, I. and W. Ziegler (2004) Syllable frequency and syllable structure in apraxia of speech. *Brain and Language* 88: 148–159.
- Alajouanine, T., A. Ombredane and M. Durand (1939) *Le Syndrome de la Désintégration Phonétique dans l'Aphasie*. Paris: Masson.
- Albareda-Castellot, B., F. Pons and N. Sebastián-Gallés (in press) The acquisition of phonetic categories in bilingual infants: New data from a new paradigm. *Developmental Science*.
- Alderete, J. (1999) Morphologically Governed Accent in Optimality Theory. Ph.D. dissertation, University of Massachusetts, Amherst. [New York: Routledge, 2001].
- Alderete, J., A. Brasoveanu, N. Merchant, A. Prince and B. Tesar (2005) Contrast analysis aids in the learning of phonological underlying forms. *The Proceedings of West Coast Conference on Formal Linguistics* 24: 34–42.
- Alexiadou, A. and E. Anagnostopoulou (1998) Parametrizing AGR: Word-order, V-movement and EPP checking. *Natural Language and Linguistic Theory* 16: 491–539.
- Allen, W.S. (1968/1999) *Vox Graeca*. Cambridge: Cambridge University Press.
- Altmann, G.T.M. and Y. Kamide (2007) The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language* 57: 502–518.
- Altmann, H. (2006) The Perception and Production of Second Language Stress: A Cross-Linguistic Experimental Study. Ph.D. dissertation, University of Delaware.
- Ambrasas, V. (ed.) (1997) *Lithuanian Grammar*. Vilnius: Institute of the Lithuanian Language.
- Amimi, A. (1996) Les deux formes nominales CəCC et CCəC en arabe marocain. Ms., Université Paris 8.
- Anani, M. (1989) Incorrect stress placement in the case of Arab learners of English. *International Review of Applied Linguistics* 27: 15–22.
- Anderson, J.M. and C. Jones (1974) Three theses concerning phonological representations. *Journal of Linguistics* 10: 1–26.
- Anderson, J.M. and C.J. Ewen (1987) *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.
- Anderson, J.M., C.J. Ewen and J. Staun (1985) Phonological structure: segmental, suprasegmental and extrasegmental. *Phonology Yearbook* 2: 203–224.
- Anderson, J.M., J.L. Morgan and K.S. White (2003) A statistical basis for speech sound discrimination. *Language and Speech* 46: 155–182.
- Anderson, S.R. (1976) Nasal consonants and the internal structure of segments. *Language* 52: 326–344.
- Anderson, S.R. (1981) Why phonology Isn't 'Natural'. *Linguistic Inquiry* 12: 493–539.

- Anderson, S.R. (1985) *Phonology in the Twentieth Century*. Chicago, IL: University of Chicago Press.
- Anttila, A. (2002) Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20: 1–42.
- Anttila, A., V. Fong, Š. Beňuš, and J. Nycz (2008) Variation and opacity in Singapore English consonant clusters. *Phonology* 25: 181–216.
- Aoun, Y. (1979) Is the syllable or the supersyllable a constituent? *MIT Working Papers in Linguistics* 1: 140–148.
- Apoussidou, D. (2007) *The Learnability of Metrical Phonology*. Ph.D. dissertation, University of Amsterdam.
- Apoussidou, D. and S. Nordhoff (2010) Feet in Sri Lankan Malay: No stress, please! Manuscript, University of Utrecht & Max Planck Institute for Evolutionary Anthropology.
- Archangeli, D. (1984) Underspecification in Yawelmani Phonology and Morphology. Ph.D. dissertation, Massachusetts Institute of Technology. Published in 1988 in *Outstanding Dissertations in Linguistics Series*, Garland Press, New York.
- Archangeli, D. (1988) Aspects of Underspecification Theory. *Phonology* 5: 183–207.
- Archangeli, D. and D. Pulleyblank (1994) *Grounded Phonology*. Cambridge, MA: MIT Press.
- Archibald, J. (1993) *Language Learnability and L2 Phonology*. Dordrecht: Kluwer Academic Publishers.
- Archibald, J. (1994) A formal model of learning L2 prosodic phonology. *Second Language Research* 10: 215–240.
- Archibald, J. (1997) The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics* 35: 167–181.
- Archibald, J. (1998) Metrical parameters and lexical dependency: Acquiring L2 stress, in S. Flynn, G. Martohardjono and W. O’Neil (eds) *The Generative Study of Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 279–301.
- Archibald, J. (2004) Interfaces in the prosodic hierarchy: New structures and the phonological parser. *International Journal of Bilingualism* 8: 29–50.
- Arnauld, A. and C. Lancelot (1803) *Grammaire générale et raisonnée de Port-Royal*. Paris: L’imprimerie de Munier. Commentaries by Charles Pinot Duclos.
- Aronoff, M. (2007) In the beginning was the word. *Language* 83: 803–830.
- Aronoff, M., I. Meir, C. Padden and W. Sandler (2008) The roots of linguistic organization in a new language. Special Issue of *Interaction Studies on Holophrasis, Compositionality and Protolanguage* 9: 133–153.
- Aslin, R.N. and J. Fiser (2005) Methodological challenges for understanding cognitive development in infants. *Trends in Cognitive Science* 9: 92–98.
- Aslin, R.N. and J. Mehler (2005) Near-infrared spectroscopy for functional studies of brain activity in human infants: Promise, prospects, and challenges. *Journal of Biomedical Optics* 10: 011009.
- Aslin, R.N., D.B. Pisoni, B.L. Hennessy and A.J. Perey (1981) Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development* 52: 1135–1145.
- Aslin, R.N., J.R. Saffran and E.L. Newport (1998) Computation of conditional probability statistics by human infants. *Psychological Science* 9: 321–324.
- Atkey, S. (2000) The acquisition of non-native segmental contrasts: A look at English speakers’ acquisition of Czech palatal stops, in A. James and J. Leather (eds) *Proceedings of the New Sounds 2000 Conference*. Klagenfurt: University of Klagenfurt Press, pp. 22–31.
- Atkinson, K., B. MacWhinney and C. Stoel (1969) An experiment on the recognition of babbling. *Language Behavior Research Laboratory Working Paper* 14. University of California, Berkeley.

- Auroux, S., E.F.K. Koerner, H-J. Niederehe and K. Versteegh (eds) (2006) *History of the Language Sciences* (Volume 3) (HSK 18.3). Berlin: Mouton de Gruyter.
- Avery, P. and K. Rice (1989) Segment structure and coronal underspecification. *Phonology* 6: 179–200.
- Baart, J. (2001) Acoustic Phonetics. LinguaLinks 5.0. CD-ROM. Dallas. SIL.
- Bach, E. and R.T. Harms (1972) How do languages get crazy rules?, in R. Stockwell and R. Macaulay (eds) *Linguistic Change and Generative Theory*. Bloomington, IN: Indiana University Press, pp. 1–21.
- Backley, P. (1993) Coronal: The undesirable element. *UCL Working Papers in Linguistics* 5: 301–323.
- Backley, P. and K. Nasukawa (2009a) Representing labials and velars: A single ‘dark’ element. *Phonological Studies* 12: 3–10.
- Backley, P. and K. Nasukawa (2009b) Headship as melodic strength, in K. Nasukawa and P. Backley (eds) *Strength Relations in Phonology*. Berlin and New York: Mouton de Gruyter, pp. 47–77.
- Backley, P. and T. Takahashi (1998) Element activation, in E. Cyran (ed.) *Structure and Interpretation: Studies in Phonology*. Lublin: Folium, pp. 13–40.
- Bailey, T. and K. Plunkett (2002) Phonological specificity in early words. *Cognitive Development* 17: 1265–1282.
- Baker, E., K. Croot, J. van Doorn and V.A. Reed (1999) In search of a sound theory: Investigating articulation and phonology. *Proceedings of the 1999 Speech Pathology Australia National Conference*, pp. 29–36.
- Baker, S.A., R. Michnick-Golinkoff and L.A. Petitto (2006) New insights into old puzzles from infants’ categorical discrimination of soundless phonetic units. *Language Learning and Development* 2: 147–162.
- Baker, S.A., W. Idsardi, R. Michnick-Golinkoff and L.A. Petitto (2005) The perception of handshape in American Sign Language. *Memory and Cognition* 33: 887–904.
- Ball, M.J., N. Müller and B. Rutter (2010) *Phonology for Communication Disorders*. New York and Hove: Psychology Press.
- Ballem, K. and K. Plunkett (2005) Phonological specificity in 14-month-olds. *Journal of Child Language* 32: 159–173.
- Barnes, J.A. (2001) Domain-initial strengthening and the phonetics and phonology of positional neutralization. *Proceedings of the Northeast Linguistic Society (NELS)* 32 (1): 1–20.
- Barnes, J.A. (2002) Positional Neutralization: A Phonologization Approach to Typological Patterns. Ph.D. dissertation, University of California, Berkeley.
- Barton, D. (1976) Phonemic discrimination & the knowledge of words in children under three years. *Papers and Reports on Child Language Development* 11, Stanford University, pp. 61–68.
- Barton, D. (1980) Phonemic perception in children, in G.H. Yeni-Komshian, J.F. Kavanagh and C.A. Ferguson (eds) *Child Phonology: Volume 2 – Perception*. New York: Academic Press, pp. 97–116.
- Basbøll, H. (1975) Grammatical boundaries in phonology. *Ariptic* 9: 109–135.
- Basbøll, H. (1981) On the function of boundaries in phonological rules, in D. Goyvaerts (ed.) *Phonology in the 1980s*. Ghent: Story-Scientia, pp. 245–269.
- Bastiaanse, R., D.G. Gilbers and K. van der Linde (1994) Sonorant sustainment: A single case study of a rare phonemic deficit in aphasia. *Clinical Linguistics and Phonetics* 1: 27–43.
- Battison, R. (1978) *Lexical Borrowing in American Sign Language*. Silver Spring, MD: Linstok Press.
- Baudouin de Courtenay, J. (1910 [1972]) Phonetic laws, in E. Stankiewicz (ed.) *A Baudouin de Courtenay Reader*. Bloomington, IN: Indiana University Press, pp. 260–277.

- Baudouin de Courtenay, J. (1972 [1895]) An attempt at a theory of phonetic alternations, in E. Stankiewicz (ed.) *Selected Writings of Baudouin de Courtenay*. Bloomington, IN: Indiana University Press, pp. 144–212.
- Beckman, J.N. (1997) Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 13: 283–328.
- Beckman, J.N. (1998) Positional Faithfulness. Ph.D. dissertation, University of Massachusetts, Amherst.
- Beckman, M.E. (1986) *Stress and Non-Stress Accent*. Dordrecht: Foris Publications.
- Beckman, M.E. (1996) The parsing of prosody. *Language and Cognitive Processes* 11: 17–67.
- Beckman, M.E. (2003) Input representations (Inside the mind and out), in G. Garding and M. Tsujimura (eds) *Proceedings of the 22nd West-Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press, pp. 70–94.
- Beckman, M.E. and J. Edwards (1994) Articulatory evidence for differentiating stress categories, in P.A. Keating (ed.) *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*. Cambridge: Cambridge University Press, pp. 7–33.
- Beckman, M.E. and J. Kingston (1990) Introduction, in J. Kingston and M. Beckman (eds) *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, pp. 1–16.
- Beckman, M.E. and J.B. Pierrehumbert (1986) Intonational structure in English and Japanese. *Phonology Yearbook* 3: 255–309.
- Beckman, M.E., J. Edwards and J. Fletcher (1992) Prosodic structure and tempo in a sonority model of articulatory dynamics, in G. Docherty and B. Ladd (eds) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press, pp. 68–86.
- Beddor, P.S. and R. Krakow (1999) Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *Journal of the Acoustical Society of America* 106: 2868–2887.
- Beddor, P.S., J.D. Harnsberger and S. Lindemann (2002) Language-specific patterns of vowel-to-vowel coarticulation: Acoustic structures and their perceptual correlates. *Journal of Phonetics* 30: 591–627.
- Beddor, P.S., R. Krakow and L.M. Goldstein (1986) Perceptual constraints and phonological change: A study of nasal vowel height. *Phonology Yearbook* 3: 197–217.
- Beers, M. (1995) The phonology of normally developing and language-impaired children. Amsterdam: Institute for Functional Research into Language and Language Use.
- Béland, R. and C. Paradis (1997) Principled syllabic dissolution in a primary progressive aphasia case. *Aphasiology* 11: 1171–1196.
- Béland, R., D. Caplan and J.-L. Nespoulous (1990) The role of abstract phonological representations in word production: Evidence from phonemic paraphasias. *Journal of Neurolinguistics* 5: 125–164.
- Bellugi, U. and S. Fischer (1972) A comparison of sign language and spoken language. *Cognition* 1: 173–200.
- Benua, L. (1997) Transderivational Identity: Phonological Relations between Words. Ph.D. dissertation, University of Massachusetts, Amherst.
- Berko, J. (1958) The child's learning of English morphology. *Word* 14: 150–177.
- Berkovits, R. (1993) Utterance-final lengthening and the duration of final-stop closures. *Journal of Phonetics* 21: 479–489.
- Bermúdez-Otero, R. (2004) Raising and Flapping in Canadian English: Grammar and acquisition. Handout of paper presented at the CASTL Colloquium, University of Tromsø, 2 November 2004. Available at www.bermudez-otero.com/tromsoe.pdf.
- Bermúdez-Otero, R. (forthcoming a) *Stratal Optimality Theory*. Oxford: Oxford University Press.

- Bermúdez-Otero, R. (forthcoming b) Cyclicity, in M. van Oostendorp, C. Ewen, E. Hume and K. Rice (eds) *The Blackwell Companion to Phonology*. Malden, MA: Wiley-Blackwell.
- Bermúdez-Otero, R. (to appear) *Stratal Optimality Theory: Synchronic and Diachronic Applications*. Oxford: Oxford University Press.
- Berndt, R.S., C.C. Mitchum and A.N. Haendiges (1996) Comprehension of reversible sentences in 'agrammatism': A meta-analysis. *Cognition* 58: 289–308.
- Bernhardt, B. and J.P. Stemberger (2007) Phonological impairment in children and adults, in P. de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, pp. 575–593.
- Best, C.T. (1993) Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development, in B. De Boysson-Bardies, S. De Schonen, P. Jusczyk, P. MacNeilage and J. Morton (eds) *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Dordrecht: Kluwer Academic Publishers, pp. 289–304.
- Best, C.T. (1994) The emergence of language-specific phonemic influences in infant speech perception, in J. Goodman and H.C. Nusbaum (eds) *Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. Cambridge, MA: MIT Press, pp. 167–224.
- Best, C.T. (1995) A direct realist view of cross-language speech perception, in W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium, MD: York Press, pp. 171–204.
- Best, C.T. and G.W. McRoberts (2003) Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Language and Speech* 46: 183–216.
- Best, C.T., G.W. McRoberts and E. Goodell (2001) American listeners' perception of nonnative consonant contrasts varying in perceptual assimilation to English phonology. *Journal of the Acoustical Society of America* 109: 775–794.
- Best, C.T., G.W. McRoberts and N.M. Sithole (1988) The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14: 345–360.
- Bickmore, L. (1989) Kinyambo Prosody. Ph.D. dissertation, University of California, Los Angeles.
- Bickmore, L. (1990) Branching nodes and prosodic categories: Evidence from Kinyambo, in S. Inkelas and D. Zec (eds) *The Phonology-Syntax Connection*. Chicago, IL and London: University of Chicago Press, pp. 1–17.
- Bickmore, L. (1992) Multiple phonemic stress levels in Kinyambo. *Phonology* 9: 155–198.
- Bierwisch, Manfred (1966) Regeln für die Intonation deutscher Sätze, in Manfred Bierwisch (ed.) *Untersuchungen über Akzent und Intonation im deutschen*, *Studia Grammatica*. Berlin: Akademie Verlag, pp. 99–201.
- Bird, S., J. Coleman, J. Pierrehumbert and J. Scobbie (1992) Declarative phonology, *Proceedings of the XVth International Congress of Linguists*. Laval: Presses de l'Université.
- Birdsong, D. (2005) Interpreting age effects in second language acquisition, in J.F. Kroll and A.M.B. de Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press, pp. 109–127.
- Bishop, J. (2007) Incomplete neutralization in Eastern Andalusian Spanish: Perceptual consequences of durational differences involved in s-aspiration. *Proceedings of the International Congress of Phonetic Sciences XVI*, pp. 1765–1768. [www.icphs2007.de/]
- Blaho, S. (2008) The Syntax of Phonology. A Radically Substance-free Approach. Ph.D. dissertation, University of Tromsø.
- Blaho, S., P. Bye and M. Krämer (eds) (2007) *Freedom of Analysis?* Berlin: Mouton de Gruyter.

- Blevins, J. (1995) The syllable in phonological theory, in Goldsmith (ed.) *Handbook of Phonological Theory*. Oxford: Oxford University Press, pp. 245–306.
- Blevins, J. (2002) The independent nature of phonotactic constraints: An alternative to syllable-based approaches, in C. Féry and R. van de Vijver (eds) *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press, pp. 375–403.
- Blevins, J. (2004) *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
- Blevins, J. (2006a) A theoretical synopsis of evolutionary phonology. *Theoretical Linguistics* 32: 117–166.
- Blevins, J. (2006b) Reply to commentaries. *Theoretical Linguistics* 32: 245–256.
- Bloch, B. (1948) A set of postulates for phonemic analysis. *Language* 24: 3–46.
- Blom, E. and S. Unsworth (eds) (2010) *Experimental Methods in Language Acquisition Research*. Amsterdam: John Benjamins (Language Learning and Language Teaching series).
- Bloomfield, L. (1933) *Language*. Chicago, IL: Chicago University Press.
- Bloomfield, L. (1939) Menomini morphophonemics. *Travaux du Cercle Linguistique de Prague* 8: 105–115.
- Blumstein, S.E. (1973) Some phonological implications of aphasic speech, in H. Goodglass and S. Blumstein (eds) *Psycholinguistics and Aphasia*. Baltimore, MD: Johns Hopkins University Press, pp. 123–137.
- Blumstein, S.E. (1978) Segment structure and the syllable in aphasia, in A. Bell and J.B. Hooper (eds) *Syllables and Segments*. Amsterdam: North Holland, pp. 189–200.
- Blumstein, S.E. (1991) Phonological aspects of aphasia, in M.T. Sarno (ed.) *Acquired Aphasia*. San Diego, CA: Academic Press, pp. 151–180.
- Blumstein, S.E., W.E. Cooper, E.B. Zurif and A. Caramazza (1977) The perception and production of voice-onset time in aphasia. *Neuropsychologia* 15: 371–383.
- Blumstein, S.E., W.E. Cooper, H. Goodglass, S. Statlender and J. Gottlieb (1980) Production deficits in aphasia: A voice-onset time analysis. *Brain and Language* 9: 153–170.
- Bobaljik, J.D. (2002) A-chains at the PF interface: Copies and ‘covert’ movement. *Natural Language and Linguistic Theory* 20: 197–267.
- Boersma, P. (1998) *Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives*. Ph.D. dissertation, University of Amsterdam.
- Boersma, P. (2001) Phonology-semantics interaction in OT, and its acquisition, in R. Kirchner, W. Wikeley and J. Pater (eds) *Papers in Experimental and Theoretical Linguistics* 6. Edmonton: University of Alberta, pp. 24–35.
- Boersma, P. (2007) Some listener-oriented accounts of h-aspiré in French. *Lingua* 117: 1989–2054.
- Boersma, P. (2009) Cue constraints and their interactions in phonological perception and production, in P. Boersma and S. Hamann (eds) *Phonology in Perception*. Berlin: Mouton de Gruyter, pp. 55–108.
- Boersma, P. (to appear) A programme for bidirectional phonology and phonetics and their acquisition and evolution, in A. Benz and J. Mattausch (eds) *Bidirectional Optimality Theory*.
- Boersma, P. and P. Escudero (2008) Learning to perceive a smaller L2 vowel inventory: An Optimality Theory account, in P. Avery, B.E. Dresher and K. Rice (eds) *Contrast in phonology: Theory, Perception, Acquisition*. Berlin: Mouton de Gruyter, pp. 271–301.
- Boersma, P. and S. Hamann (2008) The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217–270.
- Boersma, P. and S. Hamann (2009a) Introduction: Models of phonology in perception, in P. Boersma and S. Hamann (eds) *Phonology in Perception*. Berlin: Mouton de Gruyter, pp. 1–24.

- Boersma, P. and S. Hamann (2009b) Loanword adaptation as first-language phonological perception, in L. Wetzels and A. Calabrese (eds) *Studies in Loan Phonology*. Amsterdam: John Benjamins, pp. 11–53.
- Boersma, P. and B. Hayes (2001) Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32: 45–86.
- Boersma, P. and J. Pater (2008) Convergence properties of a gradual learning algorithm for Harmonic Grammar. [ROA#970].
- Boersma, P. and D. Weenink (2009) *Praat: Doing Phonetics by Computer* (Version 5.1.04). [www.praat.org/].
- Bolton, T.L. (1894) Rhythm. *American Journal of Psychology* 6: 145–238.
- Bond, Z.S. (1999) *Slips of the Ear: Errors in the Perception of Casual Conversation*. San Diego, CA: Academic Press.
- Bongaerts, T., C. Van Summeren, B. Planken and E. Schils (1997) Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition* 19: 447–465.
- Booij, G. (1988) Review of Marina Nespor and Irene Vogel, Prosodic Phonology. *Journal of Linguistics* 24: 515–525.
- Booij, G. (1995) *The Phonology of Dutch*. Oxford: Oxford University Press.
- Booij, G. (1996) Cliticization as prosodic integration: The case of Dutch. *The Linguistic Review* 13: 219–242.
- Borowsky, T. (1986) Topics in English and Lexical Phonology. Ph.D. dissertation, University of Massachusetts, Amherst.
- Borowsky, T. (1989) Structure preservation and the syllable coda in English. *Natural Language and Linguistic Theory* 7: 145–166.
- Bortfeld, H., E. Wruck and D. Boas (2007) Assessing infants' cortical response to speech using near-infrared spectroscopy. *Neuroimage* 34: 407–415.
- Bortfeld, H., J. Morgan, R. Golinkoff and K. Rathbun (2005) Mommy and me: Familiar names help launch babies in speech stream segmentation. *Psychological Science* 16: 298–304.
- Bosch, L. and N. Sebastian-Galles (2003) Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech* 46 (2–3): 217–243.
- Bošković, Ž. (1995) Particle movement and second position cliticization in Serbo-Croatian. *Lingua* 96: 245–266.
- Bošković, Ž. (2001) *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Amsterdam, London and New York: Elsevier.
- Botinis, Antonis (1989) Stress and prosodic structure in Greek. Ph.D. dissertation, Lund University Press.
- Botma, B. (2003) On the phonological interpretation of aspirated nasals, in M. van Oostendorp and J.M. van de Weijer (eds) *The Internal Organization of Phonological Segments*. Berlin and New York: Mouton de Gruyter, pp. 255–286.
- Botma, B. (2004) Phonological Aspects of Nasality An Element-Based Dependency Approach. Ph.D. dissertation, University of Amsterdam. Utrecht: LOT Publications.
- Botma, B. (2009) Transparency in nasal harmony and the limits of reductionism, in K. Nasukawa and P. Bäckley (eds) *Strength Relations in Phonology*. Berlin and New York: Mouton de Gruyter, pp. 79–112.
- Botma, B. and N.S.H. Smith (2006) A dependency account of the fortis-lenis contrast in Cama, in J.M. van de Weijer and B. Los (eds) *Linguistics in the Netherlands 2006*. Amsterdam and Philadelphia, PA: John Benjamins, pp. 15–27.
- Botma, B. and N.S.H. Smith (2007) A dependency-based typology of nasalization and voicing phenomena, in B. Los and M. van Koppen (eds) *Linguistics in the Netherlands 2007*. Amsterdam and Philadelphia, PA: John Benjamins, pp. 36–48.

Bibliography

- Bowern, C. (2008a) *Linguistic Fieldwork: A Practical Guide*. New York: Palgrave Macmillan. [www.ruf.rice.edu/~bowern/fieldwork/].
- Bowern, C. (2008b) Fieldwork in language contact situations. Ms. Rice University Houston, Texas [www.ruf.rice.edu/~bowern/Mypapers/BowernLanguageContact.pdf].
- Boysson-Bardies, B., P. Halle, L. Sagart and C. Durand (1989) A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language* 16: 1–18.
- Boysson-Bardies, B., L. Sagart and C. Durand (1984) Discernible differences in the babbling of infants according to target language. *Journal of Child Language* 11: 1–15.
- Boysson-Bardies, B., L. Sagart, P. Halle and C. Durand (1986) Acoustic investigation of cross-linguistic variability in babbling, in Z. Lindblom (ed.) *Precursors of Early Speech*. New York: Stockton Press, pp. 113–126.
- Boysson-Bardies, B., M.M. Vihman, L. Roug-Hellichius, C. Durand, I. Landberg and F. Aro (1992) Material evidence of infant selection from the target language: A cross-linguistic phonetic study, in C.A. Ferguson, L. Menn, C. Stoel-Gammon (eds) *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press, pp. 369–392.
- Bradlow, A.R. and T. Bent (2002) The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America* 112: 272–284.
- Bradlow, A.R. and T. Bent (2008) Perceptual adaptation to non-native speech. *Cognition* 106: 707–729.
- Bradlow, A.R., R. Akahane-Yamada, D.B. Pisoni and Y. Tohkura (1999) Training Japanese listeners to identify English /r/ and /l/: long term retention of learning perception and production. *Perception and Psychophysics* 61: 977–985.
- Bradlow, A.R., D.B. Pisoni, R. Akahane-Yamada and Y. Tohkura (1997) Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America* 101: 2299–2310.
- Brame, M.K. (1974) The cycle in phonology: Stress in Palestinian, Maltese, and Spanish. *Linguistic Inquiry* 5: 39–60.
- Brennan, M. (1990) *Word Formation in British Sign Language*. Stockholm: University of Stockholm.
- Brentari, D. (1998) *A Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press.
- Brentari, D. (2002) Modality differences in sign language phonology and morphophonemics, in R. Meier, K. Cormier and D. Quinto-Pozos (eds) *Modality in Language and Linguistic Theory*. Cambridge: Cambridge University Press, pp. 35–64.
- Brien, D. (ed.) (1992) *Dictionary of British Sign Language*. London: Faber and Faber.
- Brière, E. (1966) An investigation of phonological interference. *Language* 42: 768–796.
- Brière, E. (1968) *A Psycholinguistic Study of Phonological Interference*. The Hague: Mouton de Gruyter.
- Brockhaus, W. (1991) Final devoicing and neutralization. *UCL Working Papers in Linguistics* 3: 303–330.
- Brockhaus, W. (1995) *Final Devoicing in the Phonology of German*. Tübingen: Niemeyer.
- Bromberger, S. and M. Halle (1989) Why phonology is different. *Linguistic Inquiry* 20: 51–70.
- Broselow, E. (1999) Stress, epenthesis and segment transformation in Selayarese loans, in S. Chang, L. Liaw and J. Ruppenhofer (eds) *Proceedings of the 25th Annual Conference of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society, pp. 211–255.
- Broselow, E. (2004a) Language contact phonology: Richness of the stimulus, poverty of the base. *Proceedings of the North-Eastern Linguistic Society* 34: 1–22.
- Broselow, E. (2004b) Unmarked structures and emergent rankings in second language phonology. *International Journal of Bilingualism* 8: 51–65.

- Broselow, E. and D. Finer (1991) Parameter setting in second language phonology and syntax. *Second Language Research* 7: 35–59.
- Broselow, E., S. Chen and C. Wang (1998) The emergence of the unmarked in second language acquisition. *Studies in Second Language Acquisition* 20: 261–280.
- Browman, C. and L.M. Goldstein (1986) Towards an articulatory phonology. *Phonology Yearbook* 3: 219–252.
- Browman, C. and L.M. Goldstein (1989) Articulatory gestures as phonological units. *Phonology* 6: 201–251.
- Browman, C. and L.M. Goldstein (1990) Tiers in articulatory phonology, with some implications for casual speech, in J. Kingston and M.E. Beckman (eds) *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, pp. 341–376.
- Browman, C. and L.M. Goldstein (1992) Articulatory phonology: An overview. *Phonetica* 49: 155–180.
- Brown, C. (1998) The role of L1 grammar in the L2 acquisition of segmental structure. *Second Language Research* 14: 136–193.
- Brown, C. (2000) The interrelation between speech perception and phonological acquisition from infant to adult, in Archibald, J. (ed.) *Second Language Acquisition and Linguistic Theory*. Malden, MA/Oxford: Blackwell, pp. 4–63.
- Brown C. and J. Matthews (1997) The role of feature geometry in the development of phonemic contrasts, in S.J. Hannahs and M. Young-Scholten (eds) *Generative Studies in the Acquisition of Phonology*. Amsterdam: John Benjamins, pp. 67–112.
- Brown, R. (1958) *Words and Things*. Glencoe, IL: Free Press.
- Bruce, G. (1977) *Swedish Word Accents in Sentence Perspective*. (Travaux de L'institut de Linguistique de Lund XII). Lund: Gleerup.
- Buchwald, A.B., B. Rapp and M. Stone (2007) Insertion of discrete phonological units: An articulatory and acoustic investigation of aphasic speech. *Language and Cognitive Processes* 22: 910–948.
- Buckingham, H.W. (1987) Review: Phonemic paraphasias and psycholinguistic production models for neologistic jargon. *Aphasiology* 1: 381–400.
- Buder, E.H. (1996) Experimental phonology with acoustic phonetic methods: Formant measures from child speech, in B. Bernhardt, J. Gilbert and D. Ingram (eds) *Proceedings of the UBC International Conference on Phonological Acquisition*, Somerville, MA: Cascadilla Press, pp. 254–265.
- Burnham, D.K. (1986) Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics* 7: 207–239.
- Burns, M.S. and G.J. Canter (1977) Phonemic behavior of aphasic patients with posterior cerebral lesions. *Brain and Language* 4: 492–507.
- Burton-Roberts, N., P. Carr and G. Docherty (eds) (2000) *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: New York: Oxford University Press.
- Butterworth, B. (1992) Disorders of phonological encoding. *Cognition* 42: 261–286.
- Bybee, J. (2001) *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Bybee, J. (2006a) *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.
- Bybee, J. (2006b) From usage to grammar: The mind's response to repetition. *Language* 82: 711–733.
- Bye, P. (2001) *Virtual Phonology. Rule Sandwiching and Multiple Opacity in North Saami*. Ph.D. dissertation, University of Tromsø.
- Byers-Heinlein, K. and J.F. Werker (2009) Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word learning heuristic. *Developmental Science* 12: 815–823.

- Byrd, D. (1996) A phase window framework for articulatory timing. *Phonology* 13: 139–169.
- Byrd, D. (2000) Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica* 57: 3–16
- Byrd, D. (2006) Relating prosody and dynamic events: Commentary on the papers by Cho, Navas and Smiljanić, in L.M. Goldstein, D.H. Whalen and C.T. Best (eds) *Laboratory phonology VIII: Varieties of Phonological Competence*. Berlin and New York: Mouton de Gruyter, pp. 549–561.
- Byrd, D. and E. Saltzman (1998) Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics* 26: 173–199.
- Byrd, D. and E. Saltzman (2003) The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31: 149–180.
- Byrd, D., A. Kaun, S. Narayanan and E. Saltzman (2000) Phrasal signatures in articulation, in M.B. Broe and J.B. Pierrehumbert (eds) *Acquisition and the Lexicon: Papers in Laboratory Phonology V*. Cambridge: Cambridge University Press, pp. 70–88.
- Byrd, D., J. Krivokapić and S. Lee (2006) How far, how long: On the temporal scope of prosodic boundary effects. *Journal of the Acoustical Society of America* 120: 1589–1599.
- Cabrera-Abreu, M. (2000) *A Phonological Model for Intonation Without Low Tone*. Bloomington, IN: Indiana University Linguistics Club.
- Cairns, C.E. and M.F. Feinstein (1982) Markedness and the theory of syllable structure. *Linguistic Inquiry* 13: 193–225.
- Calabrese, A. (2005) *Markedness and Economy in a Derivational Model of Phonology*. Berlin: Mouton de Gruyter.
- Cambier-Langeveld, T. (2000) *Temporal Marking of Accent and Boundaries*. Ph.D. dissertation, University of Amsterdam. Utrecht: LOT Publications.
- Cameron, R. (1992) *Pronominal and Null Subjects Variation in Spanish: Constraints, Diacrits, and Functional Compensation*. Ph.D. dissertation, University of Pennsylvania.
- Campbell, L. (1974) Phonological features: Problems and proposals. *Language* 50: 52–66.
- Campbell, L. (1981) Generative phonology vs. Finnish phonology: Retrospect and prospect, in D. Goyvaerts (eds) *Phonology in the 80s*. Ghent: Story-Scientia, pp. 147–182.
- Canter, G.J., J.E. Trost and M.S. Burns (1985) Contrasting speech patterns in apraxia of speech and phonemic paraphasia. *Brain and Language* 24: 204–222.
- Caplan, D. and G. Waters (1992) Issues arising regarding the nature and consequences of reproduction conduction aphasia in S.E. Kohn (ed.) *Conduction Aphasia*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 117–149.
- Caramazza, A. and E.B. Zurif (1976) Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language* 3: 572–582.
- Caramazza, A., D. Chialant, R. Capasso and G. Miceli (2000) Separable processing of consonants and vowels. *Nature* 403: 428–430.
- Carlson, K., C. Clifton Jr. and L. Frazier (2001) Prosodic boundaries in adjunct attachment. *Journal of Memory and Language* 45: 58–81.
- Carr, P. (2006) Universal grammar and syntax/phonology parallelisms. *Lingua* 116: 634–656.
- Carvalho, J.B. de (2002a) *De la syllabation en termes de contours CV*. Habilitation Thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris. [www.unice.fr/dsl/tobias.htm].
- Carvalho, J.B. de (2002b) Formally-grounded phonology: From constraint-based theories to theory-based constraints. *Studia Linguistica* 56: 227–263.
- Carvalho, J.B. de (2004) L'analogie est-elle un fait fonctionnel ou grammatical ? Le cas de l'apophonie portugaise. *Corpus* 3: 101–123.

- Carvalho, J.B. de, T. Scheer and P. Ségéral (eds) (2008) *Lenition and Fortition*, Volume 99 of *Studies in Generative Grammar*, Berlin. Mouton de Gruyter.
- Casali, R. (2000) Some guidelines for phonological descriptions. Canada Institute of Linguistics. (Revised 2006).
- Casali, R. (2008) Advanced field methods course. Canada Institute of Linguistics, Ms., SIL International
- Cassimjee, F. and C. Kisseberth (1998) Optimal Domains Theory and Bantu Tonology: A case study from Isixhosa and Shingazidja, in L.M. Hyman and C. Kisseberth (eds) *Theoretical Aspects of Bantu Tone*. Stanford, CA: CSLI Publications, pp. 265–314.
- Chafe, W.L. (1968) The ordering of phonological rules. *International Journal of American Linguistics* 24: 115–136.
- Chafe, W.L. (1977) Accent and related phenomena in the Five Nations Iroquois Languages, in L. Hyman (ed.) *Studies in Stress and Accent*, Southern California Occasional Papers in Linguistics 4. Los Angeles, CA: University of Southern California, p. 76.
- Chambers, J.K. (1973) Canadian raising. *Canadian Journal of Linguistics* 18: 113–135.
- Chambers, J.K. (1978) Dakota accent, in E. Cook and J. Kaye (eds) *Linguistic Studies of Native Canada*. Vancouver: University of British Columbia Press, pp. 3–18.
- Charette, M. (1989) The minimality condition in phonology. *Journal of Linguistics* 25: 159–187.
- Charette, M. (1991) *Conditions on Phonological Government*. Cambridge: Cambridge University Press.
- Charette, M. (1992) Mongolian and Polish meet government licensing. *SOAS Working Papers in Linguistics and Phonetics* 2: 275–292.
- Charette, M. and A. Göksel (1994) Vowel harmony and switching in Turkic languages. *SOAS Working Papers in Linguistics and Phonetics* 4: 31–52. Also in K. Henryk and B. Szymanek (eds) (1996) *A Festschrift for Edmund Gussmann*. Lublin: University Press of the Catholic University of Lublin, pp. 29–56.
- Charette, M. and A. Göksel (1996) Licensing constraints and vowel harmony in Turkic languages. *SOAS Working Papers in Linguistics and Phonetics* 6: 1–25. Also in Cyran, E. (ed.) (1998) *Structure and Interpretation. Studies in Phonology*. Lublin: Foliom, pp. 65–88.
- Charles-Luce, J. (1993) The effects of semantic context on voicing neutralization. *Phonetica* 50: 28–43.
- Charles-Luce, J. and P.A. Luce (1990) Similarity neighborhoods of words in young children's lexicons. *Journal of Child Language* 17: 205–215.
- Charles-Luce, J. and P.A. Luce (1995) An examination of similarity neighbourhoods in young children's receptive vocabularies. *Journal of Child Language* 22: 727–735.
- Chen, M.Y. (1985) The syntax of phonology: Xiamen tone sandhi. Ms., University of California at San Diego.
- Chen, M.Y. (1990) What must phonology know about syntax, in S. Inkelas and D. Zec (eds) *The Phonology-Syntax Connection*. Chicago, IL and London: University of Chicago Press, pp. 19–46.
- Cheng, L.L.S. (1987) Derived domains and Mandarin third tone sandhi. *Chicago Linguistic Society* 23: 16–29.
- Cho, T. (2001) Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica* 58: 129–162.
- Cho, T. (2004) Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics* 32: 141–176.
- Cho, T. (2005) Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a,i/ in English. *Journal of the Acoustical Society of America* 117: 3867–3878.

- Cho, T. (2006) Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English, in L.M. Goldstein, D.H. Whalen and C.T. Best (eds) *Laboratory Phonology VIII: Varieties of Phonological Competence*. Berlin and New York: Mouton de Gruyter, pp. 519–548.
- Cho, T. (2008) Prosodic strengthening in transboundary V-to-V lingual movement in American English. *Phonetica* 65: 45–61.
- Cho, T. and S. Jun (2000) Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society* 36: 31–44.
- Cho, T. and P.A. Keating (2001) Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics* 29: 155–190.
- Cho, T. and P.A. Keating (2009) Effects of initial position versus prominence in English. *Journal of Phonetics* 37: 466–485.
- Cho, T. and P. Ladefoged (1999) Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27: 207–229.
- Cho, T. and J. McQueen (2005) Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics* 33: 121–157.
- Cho, T., S. Jun and P. Ladefoged (2002) Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30: 193–228.
- Cho, T., J. McQueen and E. Cox (2007) Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* 35: 210–243.
- Cho, Y.-M.Y. (1990) Syntax and phrasing in Korean, in S. Inkelas and D. Zec (eds) *The Phonology-Syntax Connection*. Chicago, IL and London: University of Chicago Press, pp. 47–62.
- Cholin, J. (2008) The mental syllabary in speech production: An integration of different approaches and domains. *Aphasiology* 22: 1127–1141.
- Cholin, J., W.J.M. Levelt and N.O. Schiller (2006) Effects of syllable frequency in speech production. *Cognition* 99: 205–235.
- Chomsky, N. (1951) *Morphophonemics of Modern Hebrew*. New York: Garland Press. MA dissertation, University of Pennsylvania.
- Chomsky, N. (1964) *Current Issues in Linguistic Theory*. The Hague: Mouton de Gruyter.
- Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1970) Remarks on nominalization, in R.A. Jacobs and P.S. Rosenbaum (eds) *English Transformational Grammar*. Waltham, MA: Ginn and Co., pp. 184–221.
- Chomsky, N. (1972) *Language and Mind*. New York: Harcourt Brace Jovanovich.
- Chomsky, N. (1973) Conditions on transformations, in S. Anderson and P. Kiparsky (eds) *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston, pp. 232–286.
- Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris Publications.
- Chomsky, N. (1993) A Minimalist Program for Linguistic Theory, in K. Hale and S.J. Keyser (eds) *The View from Building 20*. Cambridge, MA: MIT Press, pp. 1–52.
- Chomsky, N. (1995) *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2000) Minimalist inquiries: The framework, in R. Martin, D. Michaels and J. Uriagereka (eds) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge MA: MIT Press, pp. 89–115.
- Chomsky, N. (2001) Derivation by phase, in M. Kenstowicz (ed.) *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press, pp. 1–52.
- Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*. New York: Harper and Row.
- Chomsky, N. and H. Lasnik (1993) The theory of Principles and Parameters, in J. Jacobs, A. von Stechow, W. Sternefeld and T. Vennemann (eds) *Syntax: An International Handbook of Contemporary Research*. Berlin: Mouton de Gruyter, pp. 506–569.

- Chomsky, N., M. Halle and F. Lukoff (1956) On accent and juncture in English, in M. Halle, H.G. Lunt, H. MacLean and C.H. van Schooneveld (eds) *For Roman Jakobson. Essays on the Occasion of his Sixtieth Birthday, 11 October 1956*, The Hague: Mouton de Gruyter, 65–80.
- Christman, S.S. (1992) Abstruse neologism formation: Parallel processing revisited. *Clinical Linguistics and Phonetics* 6: 65–76.
- Christophe, A., S. Peperkamp, C. Pallier, E. Block and J. Mehler (2004) Phonological phrase boundaries constrain lexical access: I. adult data. *Journal of Memory and Language* 51: 523–547.
- Christy, C. (1983) *Uniformitarianism in Linguistics*. Amsterdam and Philadelphia, PA: John Benjamins.
- Chun, D.M. (2002) *Discourse Intonation in L2: From Theory and Research to Practice*. Amsterdam: John Benjamins.
- Clark, J., C. Yallop and J. Fletcher (2007) *An Introduction to Phonetics and Phonology* (Third edition). Oxford: Blackwell.
- Clark, R. and I. Roberts (1993) A computational model of language learnability and language change. *Linguistic Inquiry* 24: 299–345.
- Clements, G.N. (1977) The autosegmental treatment of vowel harmony, in W.U. Dressler and O. Pfeiffer (eds) *Phonologica 1976*. Innsbruck: Universität Innsbruck, pp. 111–119.
- Clements, G.N. (1985) The geometry of phonological features. *Phonology Yearbook* 2: 225–252.
- Clements, G.N. (1990) The role of sonority in core syllabification, in J. Kingston and M.E. Beckman (eds) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press, pp. 283–333.
- Clements, G.N. (1991) Vowel height assimilation in Bantu languages, in *Proceedings of the Special Session on African Language Structures (BLS 17S)*, pp. 25–64.
- Clements, G.N. (2000a) Some antecedents of nonlinear phonology. *Folia Linguistica* 34: 29–55.
- Clements, G.N. (2000b) In defense of serialism. *The Linguistic Review* 17: 181–197.
- Clements, G.N. (2001) Representational economy in constraint-based phonology, in T. Hall (ed.) *Distinctive Feature Theory*. Berlin and New York: Mouton de Gruyter, pp. 71–146.
- Clements, G.N. and S.R. Hertz (1991) Nonlinear phonology and acoustic interpretation. *Proceedings of the XIIth International Congress of Phonetic Sciences (Vol. 1/5)*. Provence: Université de Provence, pp. 364–373.
- Clements, G.N. and E. Hume (1995) The internal organization of speech sounds, in J.A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Oxford: Blackwell, pp. 245–306.
- Clements, G.N. and S.J. Keyser (1983) *CV Phonology: A Generative Theory of the Syllable*. Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.
- Clements, G.N. and A. Rialland (2007) Africa as a phonological area, in B. Heine and D. Nurse (eds) *A Linguistic Geography of Africa*. Cambridge: Cambridge University Press, pp. 36–85.
- Clumeck, H. (1976) Patterns of soft palate movements in six languages. *Journal of Phonetics* 4: 337–351.
- Coady, J.A. and R.N. Aslin (2003) Phonological neighbourhoods in the developing lexicon. *Journal of Child Language* 30: 441–469.
- Coady, J.A. and J.L. Evans (2008) Uses and interpretations of non-word repetition tasks in children with and without specific language impairment (SLI). *International Journal of Language and Communication Disorders* 43: 1–40.

- Code, C. (1998) Models, theories and heuristics in apraxia of speech. *Clinical Linguistics and Phonetics* 12: 47–65.
- Code, C. and M.J. Ball (1988) Apraxia of speech: The case for a cognitive phonetics, in M.J. Ball (ed.) *Theoretical Linguistics and Disordered Language*. London and Sydney: Croom Helm, pp. 152–167.
- Cohn, A. (1990) Phonetics and Phonological Rules of Nasalization. Ph.D. dissertation, UCLA [Also published as *UCLA Working Papers in Phonetics* 76].
- Cohn, A. (1993) Nasalization in English: phonology or phonetics? *Phonology* 10: 43–81.
- Cohn, A. (1998) The phonetics-phonology interface revisited: Where's phonetics?, in Amanda R. Doran, Tivoli Majors, Claude E. Mauk, and Nisha Merchant Goss (eds) *Texas Linguistics Society 1998 Conference Proceedings*. Austin, TX: Department of Linguistics, University of Texas, pp. 25–40.
- Cohn, A. (2006) Is there gradient phonology?, in G. Fanselow, C. Fery, R. Vogel and M. Schlesewsky (eds) *Gradience in Grammar: Generative Perspectives*. Oxford: Oxford University Press, pp. 25–44.
- Cohn, A. (2007) Phonetics in phonology and phonology in phonetics. *Working Papers of the Cornell Phonetics Laboratory* 16: 1–31.
- Cohn, A. (in press) Laboratory phonology: Past successes and current questions, challenges, and goals, in *Laboratory Phonology X*, Berlin: Mouton de Gruyter.
- Cole, J. (1995) The Cycle in phonology, in J. Goldsmith (eds) *The Handbook of Phonological Theory*. Oxford: Blackwell, pp. 70–113.
- Cole, J. and C. Kisseberth (1994) An Optimal Domains Theory of Harmony. *Studies in the Linguistic Sciences* 24: 101–114.
- Cole, J., H. Kim, H. Choi and M. Hasegawa-Johnson (2007) Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from radio news speech. *Journal of Phonetics* 35: 180–209.
- Coleman, J. (1995) Declarative lexical phonology, in J. Durand and F. Katamba (eds) *Frontiers of Phonology: Atoms, Structures, Derivations*. London: Longman, pp. 333–382.
- Coleman, J. (1999) The nature of vocoids associated with syllabic consonants in Tashlhiyt Berber, in *Proceedings of the 14th International Congress of Phonetic Sciences*. San Francisco, CA: pp. 735–738.
- Comrie, B. (1976) Irregular stress in Polish and Macedonian. *International Review of Slavic Linguistics* 1: 227–240.
- Conboy, B.T., J. Sommerville and P.K. Kuhl (2008) Cognitive control factors in speech perception at 11 months. *Developmental Psychology* 44 (5): 1505–12.
- Cooper, A. (1991) Glottal Gestures and Aspiration in English. Ph.D. dissertation, Yale University.
- Corver, N. and J. Nunes (2007) *The Copy-Theory of Movement*. Amsterdam: John Benjamins.
- Coulter, G. (1982) On the nature of ASL as a monosyllabic language. Paper presented at the Annual Meeting of the Linguistic Society of America, San Diego.
- Couper-Kuhlen, E. and M. Selting (eds) (1996) *Prosody in Conversation: Interactional Studies*. Cambridge: Cambridge University Press.
- Cowper, E.A. and K.D. Rice (1987) Are phonosyntactic rules necessary? *Yearbook of Phonology* 4: 185–194.
- Crasborn, O. (2006) Non-manual structures in sign language, in K. Brown (ed.) *Encyclopedia of Language and Linguistics* (Second edition, Volume 8). Oxford: Elsevier, pp. 668–672.
- Crowley, T. (2007) *Field Linguistics: A Beginner's Guide*. Oxford: Oxford University Press.
- Currie Hall, K. (2005) Defining Phonological Rules over Lexical Neighbourhoods: Evidence from Canadian Raising, in J. Alderete, C. Han and A. Kochetov (eds)

- Proceedings of the Twenty-Fourth West Coast Conference on Formal Linguistics (WCCFL 24)*, pp. 191–199.
- Currie Hall, K. (2009) A Probabilistic Model of Phonological Relationships from Contrast to Allophony. Ph.D. dissertation, The Ohio State University.
- Curtin, S. and D.G. Hufnagle (2009) Prelinguistic speech perception, in E. Bavin (ed.) *The Cambridge University Press Handbook of Child Language*. Cambridge: Cambridge University Press, pp. 107–124.
- Curtin, S., C. Fennell and P. Escudero (2009) Weighting of acoustic cues explains patterns of word-object associative learning. *Developmental Science*, 12: 725–731.
- Curtin, S., H. Goad and J. Pater (1998) Phonological transfer and levels of representation: The perceptual acquisition of Thai voice and aspiration by English and French speakers. *Second Language Research* 14: 389–405.
- Cutler, A. (1976) Phoneme monitoring as a function of preceding intonation contour. *Perception and Psychophysics* 20: 55–60.
- Cutler, A. (ed.) (1982) *Slips of the tongue and language production*. Berlin: Mouton.
- Cutler, A. and D. Norris (1979) Monitoring sentence comprehension, in W.E. Cooper and E.C.T. Walker (eds) *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 113–134.
- Cutler, A. and D. Norris (1988) The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 14: 113–121.
- Cutler, A. and D. Norris (1988) The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human* 14: 113–121.
- Cutler, A. and D. Pasveer (2006) Explaining cross-linguistic differences in effects of lexical stress on spoken-word recognition, in R. Hoffman and H. Mixdorff (eds) *Speech Prosody, 3rd International Conference, Dresden, May 2–5, 2006, Proceedings*. Dresden: TUD Press, pp. 237–400.
- Cutler, A. and S. Butterfield (1992) Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language* 31: 218–236.
- Cutler, A., A. Weber and T. Otake (2006) Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics* 34: 269–284.
- Cutting, J.E. and B.S. Rosner (1974) Categories and boundaries in speech and music. *Perception and Psychophysics* 16: 564–571.
- Cutting, J.E., B.S. Rosner and C.R. Foard (1976) Perceptual categories for music-like sounds: Implications for theories of speech perception. *Quarterly Journal of Experimental Psychology* 28: 361–378.
- Cyran, E. (1997) *Resonance Elements in Phonology: A Study in Munster Irish*. Lublin: Folium.
- Cyran, E. (2003) *Complexity Scales and Licensing Strength in Phonology*. Lublin: KUL.
- Cyran, E. (2010) *Complexity Scales and Licensing Strength in Phonology*. Berlin: Walter de Gruyter.
- Cyran, E. and E. Gussmann (1999) Consonantal clusters and governing relations: Polish initial consonant sequences, in H. van der Hulst and N. Ritter (eds) *The Syllable. Views and Facts*. Berlin: Mouton de Gruyter, pp. 219–247.
- Daelemans, W., S. Gillis and G. Durieux (1994) The acquisition of stress: a data-oriented approach. *Computational Linguistics* 20: 421–451.
- Dale, P.S., E. Simonoff, D.V.M. Bishop, T.C. Eley, B. Oliver and T.S. Price (1998) Genetic influence on language delay in two-year-old children. *Nature Neuro-Science* 1: 324–328.
- Davenport, M. (1994) The characterization of nasality in dependency phonology, in H. van der Hulst and J.M. van de Weijer (eds) *Leiden in Last (HIL Phonology Papers I)*. The Hague: Holland Academic Graphics, pp. 89–103.

- Davidson, L. (2006) Phonology, phonetics, or frequency : Influences on the production of non-native sequences. *Journal of Phonetics* 34: 104–137.
- Davis, M.H., W.D. Marslen-Wilson and M.G. Gaskell (2002) Leading up the lexical garden-path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 28: 218–244.
- Davis, S. (1988) Syllable onsets as a factor in stress rules. *Phonology* 5: 1–19.
- De Boer, B. (2001) *The Origins of Vowel Systems*. Oxford: Oxford University Press.
- De Jong, K. (1991) The Oral Articulation of English Stress Accent. Ph.D. dissertation, Ohio State University.
- De Jong, K. (1995) The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97: 491–504.
- DeCasper, A.J., J.P. Lecanuet, R. Maugais, C. Granier-Deferre and M.C. Busnel (1994) Fetal reactions to recurrent maternal speech. *Infant Behavioral & Development* 17: 159–164.
- DeCasper, A.J. and W.P. Fifer (1980) Of human bonding: Newborns prefer their mothers voices. *Science* 208: 1174–1176.
- Dehaene-Lambertz, G., E. Dupoux and A. Gout (2000) Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience* 12: 635–647.
- DeKeyser, R. and J. Larson-Hall (2005) What does the critical period really mean? in J.F. Kroll and A.M.B. de Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press, pp. 88–108.
- Delattre, P. (1940) La force d'articulation consonantique en français. *The French Review* 14: 220–232.
- Dell, F. (1984) L'accentuation des phrases en français, in F. Dell, D. Hirst and J.-R. Vergnaud (eds) *Forme Sonore du Langage*. Paris: Hermann, pp. 65–122.
- Dell, F. (1985) *Les Règles et les Sons*. Paris: Hermann.
- Dell, F. and M. Elmedlaoui (1985) Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7: 105–130.
- Dell, G., M.F. Schwartz, N. Martin, E.M. Saffran and D.A. Gagnon (1997) Lexical access in aphasic and nonaphasic speakers. *Psychological Review* 104: 801–838.
- Dell, G.S. (1986) A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93: 283–321.
- Dell, G.S. (1988) The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language* 27: 124–142.
- Demuth, K. (2008) Exploiting corpora for language acquisition research, in H. Behrens (ed.) *Corpora in Language Acquisition Research: Finding Structure in Data*. Trends in Language Acquisition Research (TILAR) Series 6, Amsterdam: John Benjamins, pp. 199–205.
- Demuth, K., J. Culbertson and J. Alter (2006) Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language and Speech* 49: 137–174.
- Den Ouden, D.B. (2002) Phonology in Aphasia: Syllables and Segments in Level-Specific Deficits. Ph.D. dissertation, University of Groningen.
- Den Ouden, D.B. (2004) Multi-level OT: An argument from speech pathology, in L. Cornips and J. Doetjes (eds) *Linguistics in the Netherlands 2004*. Amsterdam and Philadelphia, PA: John Benjamins, pp. 146–157.
- Den Ouden, D.B. and R. Bastiaanse (2005) Phonological encoding and conduction aphasia in R.J. Hartsuiker, R. Bastiaanse, A. Postma and N.K. Wijnen (eds) *Phonological Encoding and Monitoring in Normal and Pathological Speech*. London: Psychology Press, pp. 86–101.
- Devine, A.M. and L.D. Stephens (1994) *The Prosody of Greek Speech*. New York and Oxford: Oxford University Press.

- Dienes, P. and P. Szigetvári (1999) Repartitioning the skeleton: VC Phonology. Ms., Eötvös Loránd University, Budapest.
- Dietrich, C., D. Swingley and J.F. Werker (2007) Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences* 104: 16027–16031.
- Dinnsen, D. (1974) Constraints on global rules in phonology. *Language* 50: 29–51.
- Dinnsen, D. (1979) Atomic phonology, in D. Dinnsen (ed.) *Current Approaches to Phonological Theory*. Bloomington, IN: Indiana University Press, pp. 31–49.
- Dinnsen, D. (1980) Phonological rules and phonetic explanation. *Journal of Linguistics* 16: 171–191.
- Dinnsen, D. (1985) A re-examination of phonological neutralization. *Journal of Linguistics* 21: 265–279.
- Dinnsen, D. and J. Charles-Luce (1984) Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics* 12: 49–60.
- Dinsmore, J. (ed.) (1992) *The Symbolic and Connectionist Paradigms: Closing the Gap*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Docherty, G. (1992) *The timing of British English obstruents*. Berlin: Foris.
- Docherty, G. and P. Foulkes (1999) Instrumental phonetics and phonological variation: Case studies from Newcastle upon Tyne and Derby, in P. Foulkes and G. Docherty (eds) *Urban Voices: Accent Studies in the British Isles*. London: Arnold, pp. 47–71.
- Docherty, G., P. Foulkes, J. Tillotson and D.J.L. Watt (2006) On the scope of phonological learning: Issues arising from socially structured variation, in L.M. Goldstein, D.H. Whalen and C.T. Best (eds) *Laboratory Phonology VIII*. Berlin: Mouton de Gruyter, pp. 393–421.
- Dogil, G. and B. Williams (1999) The phonetic manifestation of word stress, in H. van der Hulst (ed.) *Word Prosodic Systems in the Languages of Europe*. Berlin and New York: Mouton de Gruyter, pp. 273–334.
- Dogil, G. and J. Mayer (1998) Selective phonological impairment: A case of apraxia of speech. *Phonology* 15: 143–188.
- Dogil, G., J. Gvozdanovid and S. Kodzasov (1999) Slavic languages, in H. van der Hulst (ed.) *Word Prosodic Systems in the Languages of Europe*. Berlin and New York: Mouton de Gruyter, pp. 813–876.
- Donegan, P. and D. Stampe (1978) The syllable in phonological and prosodic structure, in A. Bell and J.B. Hooper (eds) *Syllables and Segments*. Amsterdam: North Holland, pp. 25–43.
- Donegan, P. and D. Stampe (1979) The study of natural phonology, in D. Dinnsen (ed.) *Current Approaches to Phonological Theory*. Bloomington, IN: Indiana University Press, pp. 126–173.
- Drachman, G. and A. Malikouti-Drachman (1996) Greek accentuation, in H. van der Hulst (ed.) *Word Prosodic Systems in the Languages of Europe*. Berlin and New York: Mouton de Gruyter, pp. 897–945.
- Drachman, G., R. Kager and A. Malikouti-Drachman (1995) Greek allomorphy: an optimality account, in M. Dimitrova-Vulchanova and L. Hellan (eds) *Papers from the First Conference on Formal Approaches to South Slavic Languages*, University of Trondheim Working Papers in Linguistics 28. Plovdiv: University of Trondheim, pp. 345–361.
- Dresher, E.B. (1981) Abstractness and explanation in phonology. Explanation in linguistics, in N. Hornstein and D. Lightfoot (eds) *The Logical Problem of Language Acquisition*. London and New York: Longman, pp. 76–115.
- Dresher, E.B. (1994) The prosodic basis of the Tiberian Hebrew system of accents. *Language* 70: 1–52.
- Dresher, E.B. and J. Kaye (1990) A computational learning model for metrical phonology. *Cognition* 34: 137–195.

- Dressler, W. (1974) Diachronic puzzles for natural phonology, in A. Bruck, R. Fox and M. La Galy (eds) *Papers from the Parasession on Natural Phonology*. Chicago, IL: Chicago Linguistic Society, pp. 95–102.
- Dressler, W. (1984) Explaining natural phonology. *Phonology Yearbook* 1: 29–51.
- Dupoux, E. and C. Pallier (1997) A destressing ‘deafness’ in French? *Journal of Memory and Language* 36: 406–421.
- Dupoux, E., C. Pallier, K. Kakehi and J. Mehler (2001) New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes* 16: 491–505.
- Dupoux, E., K. Kakehi, Y. Hirose, C. Pallier and J. Mehler (1999) Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25: 1568–1578.
- Dupoux, E., S. Peperkamp and N. Sebastián-Gallés (2001) A robust method to study stress ‘deafness’. *Journal of the Acoustical Society of America* 110: 1606–1618.
- Dupoux, E., N. Sebastián-Gallés, E. Navarete and S. Peperkamp (2008) Persistent stress ‘deafness’: The case of French learners of Spanish. *Cognition* 106: 682–706.
- Dziubalska-Kończak, K. (1994) Acquisitional evidence against the phonological syllable. *Papers and Studies in Contrastive Linguistics* 29: 17–30.
- Dziubalska-Kończak, K. (2002) *Beats-and-Binding Phonology*. Frankfurt am Main: Lang.
- Eckert, P. (1988) Adolescent social structure and the spread of linguistic change. *Language in Society* 17: 183–208.
- Eckman, F.R. (1977) Markedness and the contrastive analysis hypothesis. *Language Learning* 27: 315–330.
- Eckman, F.R. (1981) On predicting phonological difficulty in second language acquisition. *Studies in Second Language Acquisition* 4: 18–30.
- Eckman, F.R. (2008) Typological markedness and second language phonology, in J. Hansen and M. Zampini (eds) *Phonology and Second Language Acquisition*. Amsterdam: John Benjamins, pp. 95–116.
- Eckman, F.R. and G. Iverson (1993) Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research* 9: 234–252.
- Eckman, F.R., A. Elreyes and G.K. Iverson (2003) Some principles of second language phonology. *Second Language Research* 19: 169–208.
- Edwards, J.E. and M.E. Beckman (1988) Articulatory timing and the prosodic interpretation of syllable duration. *Phonetica* 45: 156–174.
- Edwards, J.E. and M.E. Beckman (2008) Methodological questions in studying consonant acquisition. *Clinical Linguistics and Phonetics* 22: 937–956.
- Edwards, J.E., M.E. Beckman and J. Fletcher (1991) The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America* 89: 369–382.
- Eimas, P.D. (1975) Auditory and phonetic coding of the cues for speech: Discrimination of the /r-l/ distinction by young infants. *Perception and Psychophysics* 18: 341–347.
- Eimas, P.D., E.R. Siquel, P. Jusczyk and J. Vigorito (1971) Speech perception in infants. *Science* 171: 303–306.
- Ellison, M.T. (2000) The universal constraint set: convention, not fact, in J. Dekkers, F. van der Leeuw and J. van de Weijer (eds) *Optimality Theory: Phonology, Syntax and Acquisition*. Oxford: Oxford University Press, pp. 524–553.
- Elordieta, G. (2007) Segmental phonology and syntactic structure, in G. Ramchand and C. Reiss (eds) *The Handbook of Linguistic Interfaces*. Oxford: Oxford University Press, pp. 125–177.
- Elordieta, G., S. Frota, P. Prieto and M. Vigarío (2003) Effects of constituent weight and syntactic branching on phrasing decisions in Romance. *Proceedings of the XVth International Congress of Phonetic Sciences*. Barcelona: Causal Productions.
- Emmorey, K., S. McCullough and D. Brentari (2003) Categorical perception in American Sign Language. *Language and Cognitive Processes* 18: 21–45.

- Encrevé, P. (1983) La liaison sans enchaînement. *Actes de la Recherche en Sciences Sociales* 46: 39–66.
- Encrevé, P. (1988) *La Liaison avec et sans Enchaînement: Phonologie Tridimensionnelle et Usages du Français*. Paris: Seuil.
- Encrevé, P. (1997) L'ancien et le nouveau. Quelques remarques sur la phonologie et son histoire. *Languages* 125: 100–123. (An English translation has appeared as Encrevé, P. (2000) The old and the new: Some Remarks on phonology and its history. *Folia Linguistica* 34: 57–85.).
- Engstrand, O., K. Williams and F. Lacerda (2003) Does babbling sound native? Listener responses to vocalizations produced by Swedish & American 12- and 18-month-olds. *Phonetica* 60: 17–44.
- Ericksen, D. (2002) Articulation of extreme formant patterns for emphasized vowels. *Phonetica* 59: 134–149.
- Ernestus, M. and H.M. Baayen (2006) The functionality of incomplete neutralization in Dutch: The case of past-tense formation, in L.M. Goldstein, D.H. Whalen and C.T. Best (eds) *Laboratory Phonology* 8. Berlin: Mouton de Gruyter, pp. 27–49.
- Ettlinger, M. (2009) Input-Driven Opacity. Ph.D. dissertation, Northwestern University.
- Eulitz, C. and A. Lahiri (2004) Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience* 16: 577–583.
- Everett, D. and K. Everett (1984) Syllable onsets and stress placement in Pirahã, in M. Cobler, S. MacKaye and M.T. Wescoat (eds) *Proceedings of the West Coast Conference on Formal Linguistics*. Stanford, CA: The Stanford Linguistics Organization, Volume 3, pp. 105–116.
- Ewen, C.J. (1980) The characterization of glottal stricture in dependency phonology. *York Papers in Linguistics* 8: 35–47.
- Ewen, C.J. (1995) Dependency relations in phonology, in J.A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Oxford: Blackwell, pp. 570–585.
- Ewen, C.J. and H. van der Hulst (2001) *The Phonological Structure of Words: An Introduction*. Cambridge: Cambridge University Press.
- Fadiga, L., L. Craighero, G. Buccino and G. Rizzolatti (2002) Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience* 15: 399–402.
- Fallows, D. (1981) Experimental evidence for English syllabification and syllable structure. *Journal of Linguistics* 17: 309–331.
- Favreau, Y., J.-L. Nespoulous and A.R. Lecours (1990) Syllable structure and lexical frequency effects in the phonemic errors of four aphasics. *Journal of Neurolinguistics* 5: 165–187.
- Feest, S.V.H. van der (2007) Building a Phonological Lexicon: The Acquisition of the Dutch Voicing Contrast in Perception and Production. Ph.D. dissertation, Radboud University, Nijmegen.
- Feest, S.V.H. van der and P. Fikkert (2005) Segmental detail in children's early lexical representations. *Proceedings of the ISCA Workshop on Plasticity in Speech Perception*, on CD.
- Feest, S.V.H. van der and P. Fikkert (under revision). Learning to represent phonological contrasts.
- Fennell, C.T. and S.R. Waxman (in press) What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*.
- Fennell, C.T. and J.F. Werker (2003) Early word learners' ability to access phonetic detail in well-known words. *Language and Speech* 46: 245–264.
- Fennell, C.T., S.R. Waxman and A. Weisleder (2007) With referential cues, infants successfully use phonetic detail in word learning, in H. Caunt-Nulton, S. Kulatilake,

- and I-H. Woo (eds) *Proceedings of the 31st Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, pp. 206–217.
- Fenson, L., S. Pethick, C. Renda, J.L. Cox, P.S. Dale and J.S. Reznick (1997) Technical manual and user's guide for the MacArthur communicative development inventories: Short form versions. Unpublished manuscript, San Diego State University.
- Fernald, A. (1985) Four-month-old infants prefer to listen to motherese. *Infant Behaviour and Development* 8: 181–195.
- Fernald, A., R. Zangl, A.L. Portillo and V.A. Marchman (2008) Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children, in I.A. Sekerina, E.M. Fernandez and H. Clahsen (eds) *Developmental Psycholinguistics: On-Line Methods in Children's Language Processing*. Amsterdam: John Benjamins, pp. 97–135.
- Fikkert, P. (1994) On the Acquisition of Prosodic Structure. Ph.D. dissertation, Holland Academic Graphics, The Hague.
- Fikkert, P. (2007) Acquiring phonology, in P. de Lacy (ed.) *Handbook of Phonological Theory*. Cambridge: Cambridge University Press, pp. 537–554.
- Fikkert, P., C. Levelt and T.S. Zamuner (2005) Perception through production? Ms., Radboud University of Nijmegen and Leiden University.
- Firth, J.R. (1948) Sounds and prosodies. *Transactions of the Philological Society*, 127–152.
- Flege, J.E. (1987) The production of 'new' and 'similar' phones in foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47–65.
- Flege, J.E. (1990) English vowel production by Dutch talkers: More evidence for the 'similar' vs. 'new' distinction, in J. Leather and A. James (eds) *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second-Language Speech*. Amsterdam: University of Amsterdam, pp. 255–293.
- Flege, J.E. (1995) Second language speech learning: Theory, findings, and problems, in W. Strange (ed.) *Speech Perception and Linguistic Experience*. Timonium, MD: York Press, pp. 229–273.
- Flemming, E. (1995) Auditory Features in Phonology. Ph.D. dissertation, UCLA.
- Flemming, E. (1995/2002) *Auditory Representations in Phonology*. London: Routledge.
- Flemming, E. (1997) Phonetic detail in phonology: Towards a unified account of assimilation and coarticulation, in K. Suzuki and D. Elzinga (eds) *Proceedings of the 1995 Southwestern Workshop in Optimality Theory (SWOT)*. University of Arizona.
- Flemming, E. (2001) Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18: 7–44.
- Flemming, E. (2003) The relationship between coronal place and vowel backness. *Phonology* 20: 335–373.
- Flemming, E. (2004) Contrast and perceptual distinctiveness, in B. Hayes, R. Kirchner and D. Steriade (eds) *The Phonetic Bases of Markedness*. Cambridge: Cambridge University Press, pp. 232–276.
- Flemming, E. (2005) Speech perception and phonological contrast, in D.B. Pisoni and R.E. Remez (eds) *The Handbook of Speech Perception*. Oxford: Blackwell, pp. 156–181.
- Flora, J. (1974) Palauan Phonology and Morphology. Ph.D. dissertation, University of California, San Diego.
- Flynn, S. (1983) A Study of the Effects of Principal Branching Direction in Second Language Acquisition: The Generalization of a Parameter of Universal Grammar from First to Second Language Acquisition. Ph.D. dissertation, Cornell University.
- Fodor, J. (1983) *The Modularity of the Mind*. Cambridge, MA: MIT-Bradford.
- Fodor, J. and Z. Pylyshyn (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition* 28: 3–71.

- Foley, J. (1977) *Foundations of Theoretical Phonology*. Cambridge: Cambridge University Press.
- Fougeron, C. (1999) Prosodically conditioned articulatory variations: A review. *UCLA Working Papers in Phonetics* 97: 1–74.
- Fougeron, C. (2001) Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29: 109–135.
- Fougeron, C. and P.A. Keating (1997) Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 106: 3728–3740.
- Foulkes, P. and G. Docherty (2007) Phonological and prosodic variation in the English of England, in D. Britain (ed.) *Language in the British Isles*. Cambridge: Cambridge University Press, pp. 52–74.
- Foulkes, P., J.M. Scobbie, D.J.L. Watt (to appear) Sociophonetics, in W.J. Hardcastle, J. Laver and F. Gibbon (eds) *Handbook of Phonetic Sciences* (Second edition). Oxford: Blackwell.
- Fourakis, M. and R. Port (1986) Stop epenthesis in English. *Journal of Phonetics* 14: 197–221.
- Fowler, C.A. (1986), An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics* 14: 3–28.
- Fowler, C.A. (1995) Acoustic and kinematic correlates of contrastive stress accent in spoken English, in F. Bell-Berti and J.J. Raphael (eds) *Producing Speech: Contemporary Issues: For Katherine Safford Harris*. New York: AIP Publishing, pp. 355–373.
- Franks, S. (2000) Clitics at the interface: An introduction to clitic phenomena in European languages, in F. Beukema and M. den Dikken (eds) *Clitic Phenomena in European Languages*. Amsterdam: John Benjamins, pp. 1–46.
- Freeland, L.S. (1951) Language of the Sierra Miwok. Memoir 6 of *International Journal of American Linguistics*. Bloomington, IN: Indiana University.
- Friedrich, C.K., A. Lahiri and C. Eulitz (2008) Neurophysiological evidence for underspecified representations: Asymmetries with word initial variations. *Journal of Experimental Psychology: Human Perception and Performance* 34: 1545–1559.
- Frisch, S.A. (2004) Language processing and segmental OCP effects, in B. Hayes, R. Kirchner and D. Steriade (eds) *Phonetically-Based Phonology*. Cambridge: Cambridge University Press, pp. 346–371.
- Frisch, S.A. and B. Zawaydeh (2001) The psychological reality of OCP-place in Arabic. *Language* 77: 91–106.
- Frisch, S.A., J.B. Pierrehumbert and M.B. Broe (2004) Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22: 179–228.
- Fromkin, V.A. (1971) The non-anomalous nature of anomalous utterances. *Language* 47: 27–52.
- Fromkin, V.A. (1975) The interface between phonetics and phonology, *UCLA Working Papers in Phonetics* 31: 104–107.
- Fry, D. (1955) Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustic Society of America* 27: 765–768.
- Fudge, E.C. (1969) Syllables. *Journal of Linguistics* 5: 253–286.
- Ganong, W.F. (1980) Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6: 110–125.
- Garde, P. (1961) Réflexions sur les différences phonétiques entre les langues slaves. *Word* 17: 34–62.
- Gardner, H. (1985) *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.
- Gazdar, G., E. Klein, G. Pullum, and I. Sag (1985) *Generalized Phrase Structure Grammar*. Oxford: Blackwell.

- Gerfen, C. (1999) *Phonology and Phonetics in Coatzacoapan Mixtec*. Dordrecht: Kluwer Academic Publishers.
- Gerfen, C. and K. Hall (2001) Coda aspiration and incomplete neutralization in Eastern Andalusian Spanish. Ms., University of North Carolina at Chapel Hill.
- Gerken, L.A. (1994) A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language* 21: 565–584.
- Gerken, L.A. (2009) *Language Development*. San Diego, CA: Plural Publishing.
- Gervain, J., F. Macagno, S. Cogoi, M. Peña and J. Mehler (2008) The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences of the USA* 105: 14222–14227.
- Geschwind, N. (1965), Disconnexion syndromes in animals and man. *Brain*, 88, 585–644.
- Ghini, M. (1993) Ø-formation in Italian: A new proposal. *Toronto Working Papers in Linguistics* 12: 41–78.
- Giegerich, H.J. (1985) *Metrical Phonology and Phonological Structure*. Cambridge Studies in Linguistics 43. Cambridge: Cambridge University Press.
- Giegerich, H.J. (1992) *English Phonology: An Introduction*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Giegerich, H.J. (1999) *Lexical Strata in English*. Cambridge: Cambridge University Press.
- Gierut, J.A. (2001) A model of lexical diffusion in phonological acquisition. *Clinical Linguistics and Phonetics* 15: 19–22.
- Gierut, J.A. and H.L. Storkel (2002) Markedness and the grammar in lexical diffusion of fricatives. *Clinical Linguistics and Phonetics* 16: 115–134.
- Gilbers, D., K. van der Linde, R. Bastiaanse (1997) Phonological length, phonetic duration and aphasia. *Clinical Linguistics and Phonetics* 11: 411–422.
- Gluck, M.A. and G.H. Bower (1988) Evaluating an adaptive network model of human learning. *Journal of Memory and Language* 27: 166–195.
- Goedemans, R. (1996) An optimality account of onset-sensitive stress in quantity-insensitive languages. *The Linguistic Review* 13: 33–48.
- Goedemans, R. and E. van Zanten (2007) Stress and accent in Indonesian, in V. van Heuven and E. van Zanten (eds) *Prosody in Indonesian Languages*, LOT Occasional Series 9. Utrecht: LOT, pp. 35–62.
- Goldinger, S.D. (1996) Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 1166–1183.
- Goldinger, S.D. (1997) Words and voices: Perception and production in an episodic lexicon, in K. Johnson and J.W. Mullennix (eds) *Talker Variability in Speech Processing*. New York: Academic Press, pp. 33–56.
- Goldinger, S.D. (1998) Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105: 251–279.
- Goldrick, M. (1998) *Optimal Opacity: Covert Structure In Phonology*. Ph.D. dissertation, Johns Hopkins University.
- Goldrick, M. (2000) Turbid Output Representations and the Unity of Opacity, in *Proceedings of North East Linguistic Society (NELS)* 30.
- Goldsmith, J.A. (1974) An autosegmental typology of tone: And how Japanese fits in, in E. Kaisse and J. Hankamer (eds) *Proceedings of NELS 5*. Cambridge, MA: Harvard University Linguistics Department, pp. 172–182.
- Goldsmith, J.A. (1976a) *Autosegmental Phonology*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Goldsmith, J.A. (1976b) An overview of autosegmental phonology. *Linguistic Analysis* 2: 23–68.

- Goldsmith, J.A. (1979a) The aims of autosegmental phonology, in D. Dinnsen (eds) *Current Approaches to Phonological Theory*. Bloomington, IN: Indiana University Press, pp. 202–222.
- Goldsmith, J.A. (1979b) *Autosegmental phonology*. New York: Garland Press.
- Goldsmith, J.A. (1980) English as a tone language, in D. Goyvaerts (eds) *Phonology in the 1980s*. Gent: Story-Scientia, pp. 287–308.
- Goldsmith, J.A. (1990) *Autosegmental and Metrical Phonology*. Cambridge, MA and Oxford: Blackwell.
- Goldsmith, J.A. (1992) Local modeling in phonology, in S. Davis (ed.) *Connectionism: Theory and Practice*. Oxford: Oxford University Press, pp. 229–246.
- Goldsmith, J.A. (1993) Harmonic phonology, in J. Goldsmith (ed.) *The Last Phonological Rule*. Chicago, IL: University of Chicago Press, pp. 21–60.
- Goldsmith, J.A. (ed.) (1995) *The Handbook of Phonological Theory*. Cambridge: Blackwell.
- Goldsmith, J.A. (2005) Review of the legacy of Zellig Harris: Language and information into the 21st Century. *Language* 81: 719–736.
- Goldsmith, J.A. (2008) Generative phonology in the late 1940s. *Phonology* 25: 37–59.
- Goldsmith, J.A. and B. Laks (eds) (2000) The history of phonology in the twentieth century. *Folia Linguistica* 34: 1–2.
- Goldsmith, J.A. and B. Laks (forthcoming) Generative phonology: Its origins, its principles, and its successors, in L. Waugh and J. Joseph (eds) *The Cambridge History of Linguistics*. Cambridge: Cambridge University Press.
- Goldstein, L.M., D. Byrd and E. Saltzman (2006) The role of vocal tract gestural action units in understanding the evolution of phonology, in M. Arbib (ed.) *Action to Language via the Mirror Neuron System*. New York: Cambridge University Press, pp. 215–249.
- Goldstein, L.M., I. Chitoran and E.O. Selkirk (2007) Syllable structure as coupled oscillator modes: Evidence from Georgian and Tashlhiyt Berber, in W. Trouvain and W. Barry (eds) *Proceedings of the XVI International Congress of Phonetic Sciences*, pp. 241–244.
- Goldstein, L.M., M. Pouplier, L. Chen, E. Saltzman and D. Byrd (2007) Dynamic action units slip in speech production errors. *Cognition* 103: 386–412.
- Goodglass, H. and E. Kaplan (1983) *The Assessment of Aphasia and Related Disorders* (Second edition). Philadelphia, PA: Lea and Febiger.
- Gordon, M. (2007) Functionalism in phonology, in P. de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, pp. 61–77.
- Gow, D.W. (2002) Does English coronal assimilation create lexical ambiguity?. *Journal of Experimental Psychology: Human Perception and Performance* 28: 163–179.
- Gow, D.W. and P. Gordon (1995) Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance* 21: 344–359.
- Gow, D.W., J. Melvold and S. Manuel (1996) How word onsets drive lexical access and segmentation: Evidence from acoustics, phonology and processing. *Proceedings of ICSLP (CD-ROM)*, Philadelphia, PA.
- Goyvaerts, D. (1981) Introduction, in D. Goyvaerts (ed.) *Phonology in the 1980s*. Ghent: Story-Scientia, pp. 1–26.
- Graetz, P., R. de Bleser and K. Willmes (1992) *Akense Afasie Test. Nederlandstalige versie*. Lisse: Swets and Zeitlinger.
- Grammont, M. (1933) *Traité de phonétique*. Paris: Delgrave.
- Green, A.D. (2004) Opacity in Tiberian Hebrew. *ZAS Papers in Linguistics* 37: 37–70.
- Green, A.D. (2006) *Phonology limited*. Habilitationsschrift, Universität Potsdam.
- Green, A.D. (2007), *Phonology Limited*, [=Linguistics in Potsdam, 27]. Universität Potsdam. Rutgers Optimality Archive #745. Available at www.roa.rutgers.edu.

- Grohmann, K. (2003) *Prolific Domains: On the Anti-Locality of Movement Dependencies*. Amsterdam: John Benjamins.
- Guenther, F.H. (1995) Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review* 102: 594–621.
- Guenther, F.H., M. Hampson and D. Johnson (1998) A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review* 105: 611–633.
- Guerssel, M. (1990) On the syllabification pattern of Berber. Ms., Université du Québec à Montréal.
- Guillaume, P. (1971) *Imitation in Children*. Chicago, IL: University of Chicago Press (Original work published in 1926).
- Guion, S.G. (2005) Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition* 27: 503–533.
- Guion, S.G., J.J. Clark, T. Harada and R.P. Wayland (2003) Factors affecting stress placement for English non-words include syllabic structure, lexical class, and stress patterns of phonologically similar words. *Language and Speech* 46: 403–427.
- Guion, S.G., T. Harada and J.J. Clark (2004) Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition* 7: 207–226.
- Gupta, P. and D. Touretzky (1994) Connectionist models and linguistic theory: Investigations of stress systems in language. *Cognitive Science* 18 (1): 1–50.
- Gussenhoven, C. (1986) English plosive allophones and ambisyllabicity. *Gramma* 10: 119–141.
- Gussenhoven, G. and A.C.M. Rietveld (1992) Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics* 20: 283–303.
- Gussmann, E. (1980) *Studies in Abstract Phonology*. Cambridge, MA: MIT Press.
- Gussmann, E. (2007) *The Phonology of Polish*. Oxford: Oxford University Press.
- Gussmann, E. and J. Kaye (1993) Polish notes from a Dubrovnik Café: I. The yers. *SOAS Working Papers in Linguistics and Phonetics* 3: 427–462.
- Guthrie, M. (1968) Notes on Nzebi (Gabon). *Journal of African Languages* 7: 101–129.
- Guy, G. (1991) Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3: 1–22.
- Hagstrom, P. (1997) Contextual metrical invisibility, in B. Bruening, Y. Kang, and M. McGinnis (eds) *Papers at the Interface*, [=MIT Working Papers in Linguistics, 30]. Cambridge, MA: MITWPL, pp. 113–181.
- Haith, M.M., C. Hazan and G.S. Goodman (1988) Expectation and anticipation of dynamic visual events by 3.5 month-old babies. *Child Development* 59: 467–479.
- Hale, K. (1973) Deep-surface canonical disparities in relation to analysis and change: An Australian example, in T.A. Sebeok (ed.) *Diachronic, Areal, and Typological Linguistics*, [=Current Trends in Linguistics, 11]. The Hague: Mouton de Gruyter, pp. 401–458.
- Hale, M. (2000) Marshallese phonology, the phonetics-phonology interface and historical linguistics. *The Linguistic Review* 17: 241–257.
- Hale, M. and M. Kissock (2007) The phonetics-phonology interface and the acquisition of perseverant underspecification, in G. Ramchand and C. Reiss (eds) *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press, pp. 81–102.
- Hale, M. and C. Reiss (2008) *The Phonological Enterprise*. Oxford: Oxford University Press.
- Hall, N. (2003) *Gestures and Segments: Vowel Intrusion as Overlap*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Hall, T.A. (2000) Syllabically conditioned coalescence and deletion in Zoque: An optimality-theoretic approach. *Linguistics* 38: 711–738.

- Hall, T.A. (2001) Introduction: Phonological representations and phonetic implementation of distinctive features, in T. Hall (ed.) *Distinctive Feature Theory*. Berlin and New York: Mouton de Gruyter, pp. 1–40.
- Hall, T.A. (2007) Segmental Features, in P. de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, pp. 311–334.
- Hall, T.A. and S. Hamann (2010) On the cross-linguistic avoidance of rhotic plus high front vocoid sequences. *Lingua* 120: 1821–44.
- Halle, M. (1959) *The Sound Pattern of Russian*. The Hague: Mouton de Gruyter.
- Halle, M. (1978) Formal vs. functional considerations in phonology. *Studies in the Linguistic Sciences* 8: 123–134. (Reprinted in Brogyányi, B. (ed.) *Studies in Diachronic, Synchronic and Typological Linguistics: Festschrift for O. Szemerényi*. Amsterdam: John Benjamins, pp. 325–341.
- Halle, M. (1992) Phonological features, in W. Bright (ed.) *International Encyclopedia of Linguistics* (Volume 3). Oxford: Oxford University Press, pp. 207–212.
- Halle, M. (1995) Feature geometry and feature spreading. *Linguistic Inquiry* 26: 1–46.
- Halle, M. (1997) On stress and accent in Indo-European. *Language* 73 (2): 275–313.
- Halle, M. (1998) The stress of English words 1968–1998. *Linguistic Inquiry* 29: 539–568.
- Halle, M. (2005) Palatalization/velar softening: What it is and what it tells us about the nature of language. *Linguistic Inquiry* 36: 23–41.
- Halle, M. and O. Matushansky (2006) The morphophonology of Russian adjectival inflection. *Linguistic Inquiry* 37: 351–404.
- Halle, M. and J.-R. Vergnaud (1980) Three-dimensional phonology. *Journal of Linguistic Research* 1: 83–105.
- Halle, M. and J.-R. Vergnaud (1987) *An Essay on Stress*. Current Studies in Linguistics. Cambridge, MA and London: MIT Press.
- Halle, M., B. Vaux and A. Wolfe (2000) On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31: 387–444.
- Hallé, P., J. Segui, U. Frauenfelder and C. Meunier (1998) Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24: 592–608.
- Hamann, S. (2003) *The Phonetics and Phonology of Retroflexes*. Ph.D. dissertation, Utrecht Institute of Linguistics.
- Hamann, S. and A. Sennema, A. (2005) Voiced labiodental fricatives or glides - all the same to Germans?, in V. Hazan and P. Iverson (eds) *Proceedings of the Conference on Plasticity in Speech Processing*. London: University College London, pp. 164–167.
- Hammerly, H. (1982) Contrastive phonology and error analysis. *International Review of Applied Linguistics* 20: 17–32.
- Hancin-Bhatt, B. and R.M. Bhatt (1997) Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition* 19: 331–378.
- Hanlon, R.B. and J.A. Edmondson (1996) Disconnected phonology: A linguistic analysis of phonemic jargon aphasia. *Brain and Language* 55: 199–212.
- Hansen, K.C. and L.E. Hansen (1969) Pintupi phonology. *Oceanic Linguistics* 8 (2): 153–170.
- Hardy, J. (1993) Phonological learning and retention in second language acquisition, in F.R. Eckman (ed.) *Confluence: Linguistics, L2 Acquisition, Speech Pathology*. Amsterdam and Philadelphia, PA: John Benjamins, pp. 235–248.
- Hargus, S. and E. Kaisse (eds) (1993) *Studies in Lexical Phonology*. New York: Academic Press.
- Harms, R. (1968) *Introduction to Phonological Theory*. Englewood Cliffs, NJ: Prentice-Hall
- Harnsberger, J.D. (2000) A cross-language study of the identification of non-native nasal consonants varying in place of articulation. *Journal of the Acoustical Society of America* 108: 764–783.

- Harrington, J., J. Fletcher and C. Roberts (1995) Coarticulation and the accented/unaccented distinction: evidence from jaw movement data. *Journal of Phonetics* 23: 305–322.
- Harris, J. (1990) Segmental complexity and phonological government. *Phonology* 7: 255–300.
- Harris, J. (1993) Integrity of prosodic constituents and the domain of syllabification in Spanish and Catalan. *The view from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale & Samuel Keyser. Cambridge, Mass.: MIT Press, 177–193.
- Harris, J. (1994) *English Sound Structure*. Oxford: Blackwell.
- Harris, J. (1997) Licensing inheritance: An integrated theory of neutralisation. *Phonology* 14: 315–370.
- Harris, J. (1998) Phonological universals and phonological disorders, in E.G. Visch-Brink and R. Bastiaanse (eds) *Linguistic Levels in Aphasia*. San Diego, CA: Singular Publishing Group, pp. 91–117.
- Harris, J. (2007) Representation, in P. de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, pp. 119–138.
- Harris, J. and E. Gussmann (1998) Final codas: Why the west was wrong, in E. Cyran (ed.) *Structure and Interpretation: Studies in Phonology*. Lublin: Wydawnictwo Foliom, pp. 139–162.
- Harris, J. and G. Lindsey (1995) The elements of phonological representation, in J. Durand and F. Katamba (eds) *Frontiers of Phonology: Atoms, Structures, Derivations*. London and New York: Longman, pp. 34–79.
- Harris, J. and G. Lindsey (2000) Vowel patterns in mind and sound, in N. Burton-Roberts, P. Carr and G. Docherty (eds) *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press, pp. 185–205.
- Harris, Z. (1944) Simultaneous components in phonology. *Language* 20: 181–205.
- Harris, Z. (1951) *Methods in Structural Linguistics*. Edition 1960 entitled *Structural Linguistics*. Chicago, IL and London: University of Chicago Press.
- Haspelmath, M. (1999) Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18: 180–205.
- Haspelmath, M. (2006) Against markedness (and what to replace it with). *Journal of Linguistics* 42: 25–70.
- Haugen, E. (1956) The syllable in phonological description, in M. Halle, H.G. Lunt, H. MacLean and C.H. van Schooneveld (eds) *For Roman Jakobson. Essays on the Occasion of his Sixtieth Birthday, 11 October 1956*. The Hague: Mouton, pp. 213–221.
- Hawkins, S. (1992) An introduction to task dynamics, in G. Docherty and B. Ladd (eds) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press, pp. 9–25.
- Hayashi, W., C. Hsu and P.A. Keating (1999) Domain-initial strengthening in Taiwanese: A follow-up study. *UCLA Working Papers in Phonetics* 97: 152–156.
- Hayes, B. (1981) A Metrical Theory of Stress Rules. Ph.D. dissertation, Massachusetts Institute of Technology.
- Hayes, B. (1982) Extrametricality and English stress. *Linguistic Inquiry* 13: 227–276.
- Hayes, B. (1984) The phonology of rhythm in English. *Linguistic Inquiry* 15: 33–74.
- Hayes, B. (1986) Assimilation as spreading in Toba Batak. *Linguistic Inquiry* 17: 467–499.
- Hayes, B. (1989a) Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20: 253–306.
- Hayes, B. (1989b) [1984] The Prosodic hierarchy in meter, in P. Kiparsky and G. Youmans (eds) *Rhythm and Meter*. Orlando, FL: Academic Press, pp. 201–260.
- Hayes, B. (1989c) The prosodic hierarchy in meter, in P. Kiparsky and G. Youmans (eds) *Phonetics and Phonology, Vol. I: Rhythm and Meter*. New York: Academic Press, pp. 47–96.

- Hayes, B. (1995a) *Metrical Stress Theory: Principles and Case Studies*. Chicago, IL and London: University of Chicago Press.
- Hayes, B. (1995b) On what to teach undergraduates: Some changing orthodoxies in phonological theory, in I-H. Lee (ed.) *Linguistics in the Morning Calm* (Volume 3. Seoul: Hanshin, pp. 59–77.
- Hayes, B. (1996) Phonetically driven phonology: the role of Optimality Theory and inductive grounding. Ms., University of California, Los Angeles.
- Hayes, B. and C. Wilson (2008) A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39: 379–440.
- Hayes, B., R. Kirchner and D. Steriade (2004) *Phonetically Based Phonology*. Cambridge: Cambridge University Press.
- Hayes-Harb, R. and K. Masuda (2008) Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research* 24: 5–33.
- Heinz, J. (2006) Learning quantity-insensitive stress patterns via local inference, in *Proceedings of The Association for Computational Linguistics Special Interest Group in Phonology* 6 (ACL-SIGPHON 06).
- Heinz, J. (2009) On the role of locality in learning stress patterns. *Phonology* 26 (2): 303–351.
- Hellberg, S. (1978) Unnatural phonology. *Journal of Linguistics* 14: 157–177.
- Herrera, E. (1995) *Palabras, Estratos y Representaciones: Temas de Fonología Léxica en Zoque*. El Colegio de México: Centro de Estudios Lingüísticos y Literarios.
- Heugten, van M. and E.K. Johnson (2009) The robustness of infants' early word representations. *Journal of the Canadian Acoustical Association* 37: 148–149.
- Hickok, G. (2001) Functional anatomy of speech perception and speech production: Psycholinguistic implications. *Journal of Psycholinguistic Research* 30: 225–235.
- Hirschfeld, L. and S. Gelman (1994) Toward a topography of mind: an introduction to domain specificity, in L. Hirschfeld and S. Gelman (eds) *Mapping the Mind: Domain Specificity in Cognition and Culture*. Cambridge: Cambridge University Press, pp. 3–35.
- Hirsh, I.J. (1959) Auditory perception of temporal order. *Journal of the Acoustical Society of America* 31: 759–767.
- Hirsh, I.J. and C.E. Shetrick (1961) Perceived order in different sense modalities. *Journal of Experimental Psychology* 62: 423–492.
- Hirsh-Pasek, K. and R.M. Golinkoff (1996) The intermodal preferential looking paradigm: A window onto emerging language comprehension, in D. McDaniel, C. McKee and H.S. Cairns (eds) *Methods for Assessing Children's Syntax*. Cambridge, MA: MIT Press, pp. 105–124.
- Hochberg, J. (1986) Final /s/ deletion in Puerto Rican Spanish, in D. Sankoff (ed.) *Diachrony and Diversity*. New York: Academic Press, pp. 199–210.
- Hock, H.H. (1986) Compensatory lengthening: In defense of the concept 'mora'. *Folia Linguistica* 20: 431–460.
- Hockett, C.F. (1955) *A Manual of Phonology*. Baltimore, MD: Waverly Press.
- Hockett, C.F. (1958) *A Course in Modern Linguistics*. New York: Macmillan.
- Hockett, C.F. (1960) The origin of speech. *Scientific American* 203: 88–96.
- Hockett, C.F. (1961) Linguistic elements and their relations. *Language* 37: 29–53.
- Højen, A. and J.E. Flege (2006) Early learners discrimination of second language vowels. *Journal of the Acoustical Society of America* 119: 3072–3084.
- Hollich, G., K. Hirsh-Pasek and R. Golinkoff (2000) Breaking the language barrier: An emergentist coalition model of word learning. *Monographs of the Society for Research in Child Development* 65 (3).
- Holton, D., P. Mackridge and I. Philippaki-Warbuton (1997/2004) *Greek: A Comprehensive Grammar of the Modern Language*. London and New York: Routledge.

- Homae, F., H. Watanabe, T. Asakawa and G. Taga (2006) The right hemisphere of sleeping infants perceives sentential prosody. *Neuroscience Research* 54: 276–280.
- Homae, F., H. Watanabe, T. Nakano and G. Taga (2007) Prosodic processing in the developing brain. *Neuroscience Research* 59: 29–39.
- Hooper, J. (1972) The syllable in Linguistic Theory. *Language* 48: 525–540.
- Hooper, J. (1975) The archi-segment in natural generative phonology. *Language* 51: 536–560.
- Hooper, J. (1976) *An Introduction to Natural Generative Phonology*. New York: Academic Press.
- Horwood, G. (2001) Anti-faithfulness and Subtractive Morphology. Rutgers Optimality Archive #466. Available at www.roa.rutgers.edu.
- Houston, D.M. and P.W. Jusczyk (2000) The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance* 26: 1570–1582.
- Houston-Price, C., E. Mather and E. Sakkalou (2007) Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language* 34: 701–724.
- Hsu, C.-S. and S.-A. Jun (1998) Prosodic strengthening in Taiwanese: syntagmatic or paradigmatic?. *UCLA Working Papers in Phonetics* 96: 69–89.
- Hudson, C.L. and E.L. Newport (1999) Creolization: Could adults really have done it all?, in A. Greenhill, H. Littlefield and C. Tano (eds) *Proceedings of the Boston University Conference on Language Development* 23. Somerville, MA: Cascadilla Press, pp. 265–276.
- Hulst, H. van der (1985) Ambisyllabicity in Dutch, in H. Bennis and F. Beukema (eds) *Linguistics in the Netherlands 1985*. Dordrecht: Foris Publications, pp. 57–66.
- Hulst, H. van der (1995) Radical CV phonology: The categorial gesture, in J. Durand and F. Katamba (eds) *Frontiers of Phonology: Atoms, Structures, Derivations*. London and New York: Longman, pp. 80–116.
- Hulst, H. van der (1999) *Word Prosodic Systems in the Languages of Europe*. Berlin and New York: Mouton de Gruyter.
- Hulst, H. van der (2004) Phonological dialectics: A short history of generative phonology, in P. van Sterkenburg (ed.) *Linguistics Today: Facing a Greater Challenge*. Amsterdam: John Benjamins, pp. 217–242.
- Hulst, H. van der (2005) Why phonology is the same, in H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz and J. Koster (eds) *Organizing Grammar. Studies in Honor of Henk van Riemsdijk*. Berlin: Mouton de Gruyter, pp. 252–262.
- Hulst, H. van der and N. Ritter (1999) Head-driven phonology, in H. van der Hulst and N. Ritter (eds) *The Syllable: Views and Facts*. Berlin, New York: de Gruyter, pp. 113–167.
- Hulst, H. van der and N. Ritter (2000) The SPE-heritage of Optimality Theory. *The Linguistic Review* 17: 259–289.
- Hulst, H. van der and N. Ritter (2003) Levels, constraints and heads, in A.M. Di Sciullo (ed.) *Asymmetry in Grammar*. Amsterdam: John Benjamins, pp. 147–188.
- Hulst, H. van der and N. Smith (eds) (1982) *The Structure of Phonological Representations*. 2 Volumes. Dordrecht: Foris Publications.
- Humbert, H. (1995) *Phonological Segments: Their Structure and Behaviour*. Ph.D. dissertation, University of Leiden. The Hague: Holland Academic Graphics.
- Hume, E. (to appear) Deconstructing markedness: A predictability-based approach, in *Proceedings of the Berkeley Linguistics Society*.
- Hume, E. and K. Johnson (2001a) A model of the interplay of speech perception and phonology, in E. Hume and K. Johnson (eds) *The Role of Speech Perception in Phonology*. New York: Academic Press, pp. 3–26.

- Hume, E. and K. Johnson (eds) (2001b) *The Role of Speech Perception in Phonology*. San Diego, CA: Academic Press.
- Hurch, B. and R. Rhodes (eds) (1996) *Natural Phonology: The State of the Art*. Berlin: Mouton de Gruyter.
- Hyman, L.M. (1970) How concrete is phonology? *Language* 46: 58–76.
- Hyman, L.M. (1975) *Phonology: Theory and Analysis*. New York: Holt, Rinehart and Winston.
- Hyman, L.M. (1977) On the nature of linguistic stress, in Larry Hyman (ed.) *Studies in Stress and Accent*. Los Angeles, CA: University of Southern California, pp. 37–82.
- Hyman, L.M. (1985) *A Theory of Phonological Weight*. Dordrecht: Foris Publications.
- Hyman, L.M. (1999) The historical interpretation of vowel harmony in Bantu, in J-M. Hombert and L.M. Hyman (eds) *Bantu Historical Linguistics: Theoretical and Empirical Perspectives*. Stanford, CA: CSLI Publications, pp. 235–295.
- Hyman, L.M. (2001) The limits of phonetic determinism in Phonology: *NC revised, in E. Hume and K. Johnson (eds) *The Role of Speech Perception in Phonology*. New York: Academic Press, pp. 141–185.
- Hyman, L.M. (2002) Is there a right-to-left bias in vowel harmony? Paper presented at the Ninth International Phonology Meeting, November 1–3, 2002. Vienna. [www.ai.mit.edu/projects/dm/featgeom/hyman-VHdirectionality.pdf].
- Hyman, L.M. (2008) Universals in phonology. *The Linguistic Review* 25: 83–137.
- Idsardi, W.J. (1992) *The Computation of Prosody*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Idsardi, W.J. (1997) Phonological derivations and historical changes in Hebrew spirantization, in I. Roca (ed.) *Derivations and Constraints in Phonology*. Oxford: Clarendon Press, pp. 367–392.
- Idsardi, W.J. (2000) Clarifying opacity. *The Linguistic Review* 17: 337–350.
- Idsardi, W.J. (to appear) Canadian raising, opacity and rephonemicization, *Canadian Journal of Linguistics*.
- Ingleby, M. and W. Brockhaus (2002) Phonological primes: Cues and acoustic signatures, in J. Durand and B. Laks (eds) *Phonetics, Phonology, and Cognition*. Oxford: Oxford University Press, pp. 131–150.
- Ingram, J.C.L. and S-G. Park (1998) Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners. *Journal of the Acoustical Society of America* 103: 1161–1174.
- Inkelas, S. (1989) *Prosodic Constituency in the Lexicon*. Palo Alto, CA: Stanford University dissertation series.
- Inkelas, S. (1994) Exceptional stress-attracting suffixes in Turkish: Representations vs. the grammar, in H. van der Hulst, R. Kager and W. Zonneveld (eds) *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press, pp. 134–187.
- Inkelas S. and D. Zec (eds) (1990) *The Phonology-Syntax Connection*. Chicago, IL London: University of Chicago Press.
- Ioup, G. and S. Weinberger (eds) (1987) *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. Rowley, MA: Newbury House.
- Ishihara, S. (2003) *Intonation and Interface Conditions*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Ishihara, S. (2007) Major Phrase, Focus Intonation, Multiple Spell-Out (MaP, FI, MSO). *The Linguistic Review* 24: 137–167.
- Itô, J. (1986) *Syllable Theory in Prosodic Phonology*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Itô, J. (1989) A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7: 217–259.

- Itô, J. and A. Mester (1992) Weak layering and word binarity. *Linguistic Research Center Working Papers* LRC-92-09 UC Santa Cruz. Ms.
- Itô, J. and A. Mester (1995) Hierarchical alignment and binarity. Talk presented at TREND (University of California, Santa Cruz, April 1995) and GLOW (University of Tromsø, Norway, May 1995).
- Iverson, P. and P.K. Kuhl (2000) Perceptual magnet and phoneme boundary effects in speech perception: Do they arise from a common mechanism? *Perception and Psychophysics* 62: 874–886.
- Iverson, P., P.K. Kuhl, R. Akahane-Yamada, E. Diesch, Y. Tohkura, A. Kettermann and C. Siebert (2003) A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87: B47–B57. [www.ilabs.washington.edu/kuhl/pdf/Iverson_Kuhl_2003.pdf].
- Jakobson, R. (1939a) Observations sur le classement phonologique des consonnes. *Proceedings of the Third International Congress of Phonetic Sciences*, pp. 33–41. [Reprinted in Jakobson, R. (1962) *Selected Writings I: Phonological Studies*. The Hague: Mouton, pp. 273–279].
- Jakobson, R. (1939b) Observations sur le classement phonologique des consonnes, in E. Blancquaert and W. Pée (eds) *Proceedings of the Third International Congress of Phonetic Sciences*. Ghent: Laboratory of Phonetics of the University of Ghent, pp. 273–279.
- Jakobson, R. (1941) *Kindersprache, Aphasie und allgemeine Lautgesetze*. The Hague: Mouton.
- Jakobson, R. (1968) *Child Language, Aphasia and Phonological Universals*. The Hague: Mouton.
- Jakobson, R. and M. Halle (1956) *Fundamentals of Language*. The Hague: Mouton.
- Jakobson, R., G. Fant and M. Halle (1952) *Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates* (Second edition). Cambridge, MA: MIT Press.
- Jansma, B.M. and N.O. Schiller (2004) Monitoring syllable boundaries during speech production. *Brain and Language* 90: 311–317.
- Jarosz, G. (2006) Rich Lexicons and Restrictive Grammars: Maximum Likelihood Learning in Optimality Theory. Ph.D. dissertation, Johns Hopkins University, Baltimore, MD.
- Jensen, J.T. (2000) Against ambisyllabicity. *Phonology* 17: 187–235.
- Jesney, K. & A-M. Tessier (to appear) Biases in Harmonic Grammar: the road to restrictive learning. *Natural Language and Linguistic Theory*.
- Jespersen, O. (1904) *Lehrbuch der Phonetik*. Leipzig and Berlin: Teubner
- Jespersen, O. (1920) *Lehrbuch der Phonetik* (Third edition). Leipzig: B.G. Teubner.
- Jespersen, O. (1964) *Language: Its Nature, Development and Origin*. New York: Norton. (Original work published in 1922).
- Jessen, M. and K. Marasek (1995) Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German. *Phonetic AIMS* 2 (2): 143–146.
- Joanette, Y., E. Keller and A.R. Lecours (1980) Sequence of phonemic approximations in aphasia. *Brain and Language* 11: 30–44.
- Johnson, E.K. (2005) English-learning infants' representations of word-forms with iambic stress. *Infancy* 7: 95–105.
- Johnson, E.K. (2008) Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. *Journal of the Acoustical Society of America* 123: EL144–EL148.
- Johnson, E.K. and P.W. Jusczyk (2001) Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language* 44: 1–20.
- Johnson, E.K. and T.S. Zamuner (in press) Using infant and toddler testing methods in language acquisition research, in E. Blom and S. Unsworth (eds) *Experimental Methods in Language Acquisition Research*. Amsterdam: John Benjamins.

- Johnson, K. (1997) Speech perception without speaker normalization: An exemplar model, in K. Johnson and John W. Mullenmix (eds) *Talker Variability in Speech Processing*. New York: Academic Press, pp. 143–165.
- Johnson, K. (2005) Decisions and mechanisms in exemplar-based phonology, in *UC Berkeley Phonology Lab Annual Report 2005*. University of Berkeley, pp. 289–311.
- Johnson, R. (1991) Sign language, culture and community in a traditional Yucatec Maya village. *Sign Language Studies* 73: 461–474.
- Johnson, S.P., D. Amso and J.A. Slemmer (2003) Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences of the USA* 100: 10568–10573.
- Johnston, T. and A. Schembri (2007) *Australian Sign Language: An Introduction to Sign Language Linguistics*. Cambridge: Cambridge University Press.
- Jones, D. (1962) *The Phoneme* (Second edition). Cambridge: Heffer. First edition published 1950.
- Jongstra, W. (2003) Variable and stable clusters: Variation in the realisation of consonant clusters. *The Canadian Journal of Linguistics* 48: 265–288.
- Joos, M. (1942) A phonological dilemma in Canadian English. *Language* 18: 141–144.
- Joos, M., ed. (1957) *Readings in Linguistics, Vol. 1*. Washington, D.C.: American Council of Learned Societies.
- Joseph, J.E. (2000) *Limiting the Arbitrary*, [= *Studies in the History of the Language Sciences*, 96]. Amsterdam: John Benjamins.
- Jun, S.-A. (1993) *The Phonetics and Phonology of Korean Prosody*. Ph.D. dissertation, Ohio State University.
- Jun, S.-A. (1995) Asymmetrical prosodic effects on the laryngeal gesture in Korean, in B. Connell and A. Arvaniti (eds) *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*. Cambridge: Cambridge University Press, pp. 235–253.
- Jun, S.-A. (1998) The Accentual Phrase in the Korean prosodic hierarchy. *Phonology* 15: 189–226.
- Jun, S.-A. (2000) K-ToBI (Korean ToBI) Labelling Conventions (Version 3.1). [www.linguistics.ucla.edu/people/jun/ktobi/K-tobi.html].
- Jun, S.-A. (2003) The effect of phrase length and speech rate on prosodic phrasing. *Proceedings of the XVth International Congress of Phonetic Sciences*. Barcelona.
- Jun, S.-A., M.E. Beckman and H.-J. Lee (1998) Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. *UCLA Working Papers in Phonetics* 96: 43–68.
- Jusczyk, P. (1997) *The Discovery of Spoken Language*. Cambridge, MA: MIT Press.
- Jusczyk, P. and R. Aslin (1995) Infants' detection of the sound pattern of words in fluent speech. *Cognitive Psychology* 29: 1–23.
- Jusczyk, P., A. Friederici, J. Wessels, V. Svenkerud and A. Jusczyk (1993) Infants' sensitivity to the sound pattern of native language words. *Journal of Memory and Language* 32: 402–420.
- Jusczyk, P., D.M. Houston and M. Newsome (1999) The beginnings of word segmentation in English-learning infants. *Cognitive Psychology* 39: 159–207.
- Jusczyk, P., P. Luce and J. Charles-Luce (1994) Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33: 630–645.
- Jusczyk, P., D.B. Pisoni, A. Waley and J. Murray (1980) Discrimination of relative onset time of two-component tones by infants. *Journal of Acoustical Society of America* 67: 262–270.
- Kabak, B. (2003) *The Perceptual Processing of Second Language Consonant Clusters*. Ph.D. dissertation, University of Delaware.
- Kabak, B. and W. Idsardi (2007) Perceptual distortions in the adaptation of English consonant clusters: Syllable structure or consonantal contact constraints?. *Language and Speech* 50: 23–52.

- Kabak, B. and K. Maniwa (2007) L2 perception of English fricatives in clear and conversational speech: The role of phonetic similarity and L1 interference. *The Proceedings of the International Congress of Phonetic Sciences 2007*, pp. 781–784.
- Kabak, B. and I. Vogel (2001) The phonological word and stress assignment in Turkish. *Phonology* 18: 315–360.
- Kabak, B., K. Maniwa and N. Kazanina (to appear) Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French. *Laboratory Phonology 1*.
- Kager, R. (1993) Alternatives to the iambic-trochaic law. *Natural Language and Linguistic Theory* 11 (3): 381–432.
- Kager, R. (1999) *Optimality Theory*. Cambridge: Cambridge University Press.
- Kager, R. and W. Zonneveld (1999) Introduction, in R. Kager and W. Zonneveld (eds) *Phrasal Phonology*. Nijmegen: Nijmegen University Press, pp. 1–43.
- Kager, R., S.V.H. van der Feest, P. Fikkert, A.O. Kerkhoff and T.S. Zamuner (2007) Representations of [voice]: Evidence from acquisition, in J. van der Weijer, E.J. van der Torre (eds) *Voicing in Dutch*. Current Issues in Linguistic Theory 286. Amsterdam: John Benjamins, pp. 41–80.
- Kahn, D. (1976) Syllable-Based Generalizations in English Phonology. Ph.D. dissertation, Massachusetts Institute of Technology. Published by Garland Press, New York, 1980.
- Kahnemuyipour, A. (2004) The Syntax of Sentential Stress. Ph.D. dissertation, University of Toronto.
- Kahnemuyipour, A. (2005) On the nature of the syntax-phonology interface: The case of sentential stress. Paper presented at the *Workshop on Interfaces, Mediterranean Syntax Meeting*, University of the Aegean, 23–25 June 2005, Rhodes.
- Kaisse, E.M. (1983) The syntax of auxiliary reduction in English. *Language* 59: 93–122.
- Kaisse, E.M. (1985) *Connected Speech. The Interaction of Syntax and Phonology*. London and New York: Academic Press.
- Kaisse, E.M. (1990) Toward a typology of postlexical rules, in S. Inkelas and D. Zec (eds) *The Phonology-Syntax Connection*. Chicago, IL: Chicago University Press, pp. 127–143.
- Kaisse, E.M., and P.A. Shaw (1985) On the theory of lexical phonology. *Phonology Yearbook* 2: 1–30.
- Kaisse, E.M. and A. Zwicky (1987) Introduction: Syntactic influences on phonological rules. *Phonology* 4: 3–11.
- Karttunen, F. (1969) A note on morpheme structure in generative phonology. *Proceedings of the 1969 Conference on Computational linguistics*. Morristown, NJ: Association for Computational Linguistics, pp. 1–10.
- Karttunen, L. (1993) Finite state constraints, in J. Goldsmith (ed.) *The Last Phonological Rule*. Chicago, IL: Chicago University Press, pp. 173–194.
- Karvonen, D. (2008) Explaining nonfinality: evidence from Finnish, in C.B. Chang and H.J. Haynie (eds) *Proceedings of the 26th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, pp. 306–314.
- Kaufman, J., R. Gilmore and M.H. Johnson (2006) Frames of reference for anticipatory action in 4-month old infants. *Infant Behavior and Development* 29: 322–333.
- Kavitskaya, D. and P. Staroverov (2009) Fed counterfeeding and positional reference: Re-solving opacity, Ms., Rutgers University and Yale University.
- Kaye, J.D. (1974) Opacity and recoverability in phonology. *Canadian Journal of Linguistics* 19: 134–149.
- Kaye, J.D. (1985) On the syllable structure of certain West African languages, in Didier L. Goyvaerts (ed.) *African Linguistics: Essays in Honour of M.W.K. Semikenke*. Amsterdam and Philadelphia, PA: John Benjamins, Studies in the Sciences of Language, pp. 285–308.

- Kaye, J.D. (1989) *Phonology: A Cognitive View*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kaye, J.D. (1990a) 'Coda' licensing. *Phonology* 7: 301–330.
- Kaye, J.D. (1990b) Government in phonology: The case of Moroccan Arabic. *The Linguistic Review* 6: 131–159.
- Kaye, J.D. (1990c) What ever happened to dialect B?, in J. Mascaró and M. Nespó (eds) *Grammar in Progress: GLOW Essays for Henk van Riemsdijk*. Dordrecht: Foris Publications, pp. 259–263.
- Kaye, J.D. (1992) On the interaction of theories of lexical phonology and theories of phonological phenomena, in U. Dressler, H. Luschützky, O. Pfeiffer and J. Rennison (eds) *Phonologica 1988*. Cambridge: Cambridge University Press, pp. 141–155.
- Kaye, J.D. (1995) Derivations and interfaces, in J. Durand and F. Katamba (eds) *Frontiers of Phonology: Atoms, Structures, Derivations*. London and New York: Longman, pp. 289–332.
- Kaye, J.D. (1996) Do you believe in magic? The story of s+C sequences, in H. Kardela and B. Szymanek (eds) *A Festschrift for Edmund Gussmann from His Friends and Colleagues*. Lublin: The University Press of the Catholic University of Lublin, pp. 155–176.
- Kaye, J.D. (2001) Working with licensing constraints, in K. Dziubalska-Kołaczyk (ed.) *Constraints and Preferences*. Berlin and New York: Mouton de Gruyter, pp. 251–268.
- Kaye, J.D. (2005) GP, I'll have to put your flat feet on the ground, in H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz and J. Koster (eds) *Organizing Grammar. Studies in Honor of Henk van Riemsdijk*. Berlin: Mouton de Gruyter, pp. 283–288.
- Kaye, J.D. (2008) Canadian Raising, eh? Ms., University of Ljubljana.
- Kaye, J.D. and J. Lowenstamm (1981) Syllable structure and markedness theory, in A. Belletti, L. Brandi and L. Rizzi (eds) *Theory of Markedness in Generative Grammar*. Pisa: Scuola Normale Superiore, pp. 287–315.
- Kaye, J.D. and J. Lowenstamm (1984) De la syllabité, in F. Dell, D. Hirst and J.-R. Vergnaud, *Forme Sonore du Langage*. Paris: Hermann pp.123–159.
- Kaye, J.D., J. Lowenstamm and J.-R. Vergnaud (1985) The internal structure of phonological representations: A theory of charm and government. *Phonology Yearbook* 2: 305–328.
- Kaye, J.D., J. Lowenstamm and J.-R. Vergnaud (1990) Constituent structure and government in phonology. *Phonology* 7: 193–231.
- Kayne, R.S. (1984) *Connectedness and Binary Branching*. Dordrech: Foris Publications.
- Kayne, R.S. (1994) *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Kean, M.-L. (1974) The Strict Cycle in phonology. *Linguistic Inquiry* 5: 179–203.
- Keating, P.A. (1984) Phonetic and phonological representation of stop consonant voicing. *Language* 60: 286–319.
- Keating, P.A. (1985) Universal phonetics and the organization of grammars, in V.A. Fromkin (ed.) *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. Orlando, FL: Academic Press, pp. 115–132.
- Keating, P.A. (1988) The phonology-phonetics interface, in F. Newmeyer (ed.) *Linguistics: The Cambridge Survey. Volume I: Linguistic Theory: Foundations*. Cambridge: Cambridge University Press, pp. 281–302.
- Keating, P.A. (1990) The window model of coarticulation: Articulatory evidence, in J. Kingston and M.E. Beckman (eds) *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, pp. 451–470.
- Keating, P.A. (1996) The phonology-phonetics interface, in U. Kleinhenz (ed.) *Interfaces in Phonology*. Berlin: Akademie Verlag, pp. 262–278.
- Keating, P.A. (2006) Phonetic encoding of prosodic structure in J. Harrington and M. Tabain (eds) *Speech Production: Models, Phonetic Processes, and Techniques*. New York and Hove: Psychology Press, pp. 167–186.

- Keating, P.A. and S. Shattuck-Hufnagel (2002) A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics* 101: 112–156.
- Keating, P.A., T. Cho, C. Fougeron and C. Hsu (2003) Domain-initial strengthening in four languages, in J. Local, R. Ogden and R. Temple (eds) *Papers in Laboratory Phonology 6: Phonetic interpretations*. Cambridge: Cambridge University Press, pp. 145–163.
- Kehoe, M. and C. Stoel-Gammon (1997) Truncation patterns in English-speaking children's word productions. *Journal of Speech Language Hearing Research* 40: 526–541.
- Kehrein, W. (2002) *Phonological Representation and Phonetic Phrasing: Affricates and Laryngeals*. Tübingen: Niemeyer.
- Kehrein, W. and C. Golston (2004) A prosodic theory of laryngeal contrasts. *Phonology* 21: 325–357.
- Keller, F. and T. Alexopoulou (2001) Phonology competes with syntax: Experimental evidence for the interaction of word placement in the realization of information. *Cognition* 79: 301–372.
- Kemler Nelson, D., P.W. Jusczyk, D.R. Mandel, J. Myers, A.E. Turk and L.A. Gerken (1995) The headturn preference procedure for testing auditory perception. *Infant Behavior and Development* 18: 111–116.
- Kenstowicz, M. (1994) *Phonology in Generative Grammar*. Cambridge, MA and Oxford: Basil Blackwell.
- Kenstowicz, M. and M. Kidida (1987) The obligatory contour principle and Tangale phonology, in D. Odden (ed.) *Current Approaches to African Linguistics*, 4. Dordrecht: Foris, pp. 223–238.
- Kenstowicz, M., and C. Kisseberth (1970) Rule ordering and the asymmetry hypothesis, in M.A. Campbell (ed.) *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society (CLS 6)*, pp. 504–519.
- Kenstowicz, M., and C. Kisseberth (1973) Unmarked bleeding orders, in C.W. Kisseberth (ed.) *Studies in Generative Phonology*. Edmonton: Linguistic Research, Inc., pp. 1–12.
- Kenstowicz, M. and C. Kisseberth (1977) *Topics in Phonological Theory*. New York: Academic Press.
- Kenstowicz, M. and C. Kisseberth (1979) *Generative Phonology. Description and Theory*. New York: Academic Press.
- Kerkhoff, A. (2007) Acquisition of Morpho-Phonology: The Dutch Voicing Alternation. Ph.D. dissertation, Utrecht University.
- Kidida, M. (1985) Tangale Phonology: A Descriptive Analysis. Ph.D. dissertation, University of Illinois, Urbana Champaign.
- Kidida, M. (1993) *Tangale Phonology: A Descriptive Analysis*. Berlin: Dietrich Reimer Verlag.
- Kijak, A. (2009) How Stressful is L2 Stress? Ph.D. dissertation, Utrecht University.
- Kim, M-R. (1994) Acoustic Characteristics of Korean Stops and Perception of English Stop Consonants. Ph.D. dissertation, University of Wisconsin, Madison.
- Kim, M-R., P.S. Beddor and J. Horrocks (2002) The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics* 30: 77–100.
- Kim, S. (2003) Domain-initial strengthening of Korean fricatives. *Harvard Studies in Korean Linguistics* 9: 164–173.
- Kim, S. (2004) The Role of Prosodic Phrasing in Korean Word Segmentation. Ph.D. dissertation, UCLA.
- Kim, S. and T. Cho (2009) The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *Journal of the Acoustical Society of America* 125: 3373–3386.
- Kingston, J. (2006) The phonetics-phonology interface, in P. de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, pp. 401–434.

- Kingston, J. (2007) The phonetics-phonology interface, in P. De Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, pp. 435–456.
- Kingston, J. and M. Beckman (1990) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press.
- Kingston, J. and R.L. Diehl (1994) Phonetic knowledge. *Language* 70: 419–454.
- Kiparsky, P. (1968) Linguistic universals and language change, in E. Bach and R. Harms (eds) *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston, pp. 171–210.
- Kiparsky, P. (1968 [1973]) How abstract is phonology? Manuscript circulated since 1968 and published 1973 in O. Fujimura (ed.) *Three Dimensions of Linguistic Theory*. Tokyo: TEC, pp. 5–56.
- Kiparsky, P. (1971) Historical linguistics, in W.O. Dingwall (ed.) *A Survey of Linguistic Science*. College Park, MD: University of Maryland Linguistics Program, pp. 576–642.
- Kiparsky, P. (1972) Explanation in phonology, in S. Peters (ed.) *Goals in Linguistic Theory*. Englewood Cliffs, NJ: Prentice-Hall, pp. 189–227.
- Kiparsky, P. (1973a) Abstractness, opacity, and global rules, in O. Fujimura (ed.) *Three Dimensions of Linguistic Theory*. Tokyo: Taikusha, pp. 57–86.
- Kiparsky, P. (1973b) ‘Elsewhere’ in phonology, in S.R. Anderson and P. Kiparsky (eds) *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston, pp. 93–106.
- Kiparsky, P. (1974) On the evaluation measure, in A. Bruck, R. Fox and M. La Galy (eds) *Papers from the Parasession on Natural Phonology*. Chicago, IL: Chicago Linguistic Society, pp. 328–337.
- Kiparsky, P. (1976) Abstractness, opacity and global rules, in A. Koutsoudas (ed.) *The Application and Ordering of Grammatical Rules*. The Hague: Mouton de Gruyter, pp. 160–186.
- Kiparsky, P. (1979) Metrical structure assignment is cyclic. *Linguistic Inquiry* 10: 421–441.
- Kiparsky, P. (1982a) From cyclic phonology to lexical phonology, in H. van der Hulst and N. Smith (eds) *The Structure of Phonological Representations I*. Dordrecht: Foris Publications, pp. 131–175.
- Kiparsky, P. (1982b) Lexical phonology and morphology, in I-S. Yang (ed.) *Linguistics in the Morning Calm*. Hanshin, Seoul: Linguistic Society of Korea, pp. 3–91.
- Kiparsky, P. (1984) On the lexical phonology of Icelandic, in C.-C. Elert, I. Johansson, and E. Strangert (eds) *Nordic Prosody III, Ume å Studies in the Humanities*, 59. Stockholm: Almqvist och Wiksell, pp. 135–164.
- Kiparsky, P. (1985) Some consequences of lexical phonology. *Phonology Yearbook* 2: 85–138.
- Kiparsky, P. (1993) Blocking in nonderived environments, in S. Hargus and E. Kaisse (eds) *Studies in Lexical Phonology*. San Diego, CA: Academic Press, pp. 277–313.
- Kiparsky, P. (2000) Opacity and cyclicity. *The Linguistic Review* 17: 351–365.
- Kiparsky, P. (2002) Syllables and moras in Arabic, in C. Féry and R. van de Vijver (eds) *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press, pp. 147–182.
- Kiparsky, P. (2003) Accent, syllable structure, and morphology in Ancient Greek, in E. Mela Athanasopoulou (ed.) *Selected Papers from the 15th International Symposium on Theoretical and Applied Linguistics*. Thessaloniki, pp. 81–106.
- Kirchner, R. (1996) Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27: 341–350.
- Kirchner, R. (1997) Contrastiveness and faithfulness. *Phonology* 14: 83–111.
- Kirchner, R. (1998) An Effort-Based Approach to Consonant Lenition. Ph.D. dissertation, UCLA [Published by Routledge, 2001].
- Kirchner, R. (1998/2001) *An Effort-Based Approach to Consonant Lenition*. New York: Routledge.

- Kirchner, R. (1999/2004) Phonological contrast and articulatory effort, in B. Hayes, R. Kirchner and D. Steriade (eds) *Phonetically Based Phonology*. Cambridge: Cambridge University Press, pp. 313–345.
- Kirk, C. (2008) Substitution Errors in the Production of Word-Initial and Word-Final Consonant Clusters. *Journal of Speech, Language, and Hearing Research* 51: 35–48.
- Kisilevsky, B.S., S.M.J. Hains, K. Lee, X. Xie, H. Huang, H.H. Ye, K. Zhang and Z. Wang (2003) Effects of experience on fetal voice recognition. *Psychological Science* 14: 220–224.
- Kisseberth, C. (1970a) On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291–306.
- Kisseberth, C. (1970b) Vowel elision in Tonkawa and derivational constraints, in J.M. Sadock and A.L. Vanek (eds) *Studies Presented to Robert B. Lees by his Students*. Edmonton: Linguistic Research, Inc., pp. 109–137.
- Kisseberth, C. (1973a) Is Rule Ordering Necessary in Phonology? in B. Kachru (ed.) *Issues in Linguistics. Papers in Honor of Henry and Renée Kahane*. Urbana, IL: University of Illinois Press, pp. 418–441.
- Kisseberth, C. (1973b) On the alternation of vowel length in Klamath: A global rule, in M. Kenstowicz and C. Kisseberth (eds) *Issues in Phonological Theory*, [=Janua Linguarum, Series maior, 74]. The Hague: Mouton de Gruyter, 9–26.
- Klausenburger, J. (1978) Mikołaj Kruszewski's Theory of Morphophonology: An Appraisal. *Historiographica Linguistica* 5: 109–120.
- Klima, E. and U. Bellugi (1979) *The Signs of Language*. Cambridge, MA: Harvard University Press.
- Klin, A. (1991) Young autistic children's listening preferences in regard to speech: A possible characterization of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders* 21: 29–42.
- Koerner, E.F.K. (2002a) *Toward a History of American Linguistics*. London: Routledge.
- Koerner, E.F.K. (2002b) The Chomskyan 'Revolution' and its historiography, in E.F.K. Koerner (ed.) *Toward a History of American Linguistics*. London: Routledge, pp. 151–209.
- Koerner, E.F.K. (2004) *Essays in the History of Linguistics*. London: John Benjamins.
- Kohn, S.E. (1988) Phonological production deficits in aphasia, in H.A. Whitaker (ed.) *Phonological Processes and Brain Mechanisms*. New York: Springer.
- Kohn, S.E. (1989) The nature of the phonemic string deficit in conduction aphasia. *Aphasiology* 3: 209–239.
- Kooij, E. van der (2002) Phonological Categories in Sign Language of the Netherlands: The Role of Phonetic Implementation and Iconicity. Ph.D. dissertation, University of Leiden. Utrecht: LOT Publications.
- Koskenniemi, K. (1983) Two-level morphology: A general computational model for word-form recognition and production. Technical report, University of Helsinki. Publication no. 11.
- Koutsoudas, A. (ed.) (1976a) *The Application and Ordering of Grammatical Rules*. The Hague: Mouton de Gruyter.
- Koutsoudas, A. (1976b) Unordered rule hypotheses, in A. Koutsoudas (ed.) *The Application and Ordering of Grammatical Rules*. The Hague: Mouton de Gruyter, pp. 1–21.
- Koutsoudas, A. (1980) The question of rule ordering: Some common fallacies. *Journal of Linguistics* 16: 19–35.
- Koutsoudas, A., G. Sanders, and C. Noll (1974 [1971]) On the application of phonological rules. *Language* 50: 1–28. Originally distributed as mimeo by Indiana University Linguistics Club, Bloomington, IN.
- Kovács, A. and J. Mehler (2009a) Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences* 106: 6556–6560.

- Kovács, A. and J. Mehler (2009b) Flexible learning of multiple speech structures in bilingual infants. *Science* 325: 611–612.
- Kovelman, I., M.H. Shalinsky, K.S. White, S.N. Schmitt, M.S. Berens, N. Paymer and L.A. Petitto (2009) Dual Language use in sign-speech bimodal bilinguals: fNIRS brain-imaging evidence. *Brain and Language* 109: 112–123.
- Krämer, M. (2008) English schwa insertion before liquids and phonological opacity. *Proceedings of CLS 41*, Vol. 1, pp. 267–282.
- Kratzer, A. and E.O. Selkirk (2007) Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24: 93–135.
- Krauss, M. (1985) Supplementary notes on Central Siberian Yupik prosody, in M. Krauss (ed.) *Yupik Eskimo Prosodic Systems: Descriptive and Comparative Studies*. Fairbanks: Alaska Native Language Center, 47–50.
- Krentz, U. and D. Corina (2008) Preference for language in early infancy: The human language bias is not speech specific. *Developmental Science* 11: 1–9.
- Kristoffersen, G. (2000) *The Phonology of Norwegian*. Oxford: Oxford University Press.
- Krivokapić, J. (2007) Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics* 35: 162–179.
- Kruszewski, M. (1881[1995]) On sound alternation, in K. Koerner (ed.) *Writings in General Linguistics* (Amsterdam Classics in Linguistics 11). Amsterdam: John Benjamins, pp. 8–9.
- Kruszewski, M. (1881) *Über die Lautabwechslung*. Kazan: Universit ätsbuchdruckerei.
- Kruszewski, M. (1883[1995]) An outline of linguistic science, in K. Koerner (ed.) *Writings in General Linguistics* (Amsterdam Classics in Linguistics 11). Amsterdam: John Benjamins.
- Kuczaj, S.A. (1978) Children's judgements of grammatical and ungrammatical irregular past-tense verbs. *Child Development* 49: 319–326.
- Kuhl, P.K. (1991) Human adults and human infants show a perceptual magnet effect for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50: 93–107.
- Kuhl, P.K. (1993) Innate predispositions & the effects of experience in speech perception: The native language magnet theory, in B. deBoisson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage and J. Morton (eds) *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Dordrecht: Kluwer Academic publishers, pp. 259–274.
- Kuhl, P.K. (2000) A new view of language acquisition. *Proceedings of the National Academy of Science* 97: 11850–11857. [www.ilabs.washington.edu/kuhl/pdf/Kuhl_2000.pdf].
- Kuhl, P.K. (2008) Linking infant speech perception to language acquisition: Phonetic learning predicts language growth, in P. McCardle, J. Colombo and L. Freund (eds) *Infant Pathways to Language: Methods, Models, & Research Directions*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 213–243.
- Kuhl, P.K. and J.D. Miller (1975) Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar-plosive consonants. *Science* 190: 69–72.
- Kuhl, P.K. and J.D. Miller (1978) Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of Acoustical Society of America* 63: 905–917.
- Kuhl, P.K., E. Stevens, A. Hayashi, T. Deguchi, S. Kiritani and P. Iverson (2006) Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science* 9: 13–21.
- Kuhl, P.K., K.A. Williams, F. Lacerda, K.N. Stevens and B. Lindblom (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255: 606–608.

- Kuhn, T. (1962) *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.
- Kula, N.C. (1999) On the representation of NC clusters in Bemba, in R. van Bezooijen and R. Kager (eds) *Linguistics in the Netherlands 1999*. Amsterdam and Philadelphia, PA: John Benjamins, pp. 135–148.
- Kula, N.C. (2002) The Phonology of Verbal Derivation in Bemba. Ph.D. dissertation, University of Leiden. Utrecht: LOT Publications.
- Kula, N.C. (2006) Licensing saturation: Co-occurrence restrictions in structure. *Linguistic Analysis* 32 (3–4): 366–406.
- Kula, N.C. and L. Marten (1998) Aspects of nasality in Bemba. *SOAS Working Papers in Phonetics and Linguistics* 8: 191–208.
- Kurisu, K. (2001) The Phonology of Morpheme Realization. Ph.D. dissertation, University of California, Santa Cruz. Rutgers Optimality Archive #490. Available at www.roa.rutgers.edu.
- Kutsch Lojenga, C. (1996) Participatory research in linguistics. *Notes on Linguistics* 72: 13–27.
- Kuzla, C., T. Cho and M. Ernestus (2007) Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics* 35: 301–320.
- Labov, W. (1966) *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (1974) On the use of the present to explain the past. *Proceedings of the Eleventh Annual Conference of Linguists* 2: 825–851.
- Labov, W. (1994) *Principles of Linguistic Change* (Volume 1: Internal Factors). Oxford: Blackwell.
- Labov, W., M. Karen and C. Miller (1991) Near-mergers and the suspension of phonemic contrast. *Language Variation and Change* 3: 33–74.
- Labov, W., M. Yaeger and R. Steiner (1972) *A Quantitative Study of Sound Change in Progress*. Philadelphia, PA: U.S. Regional Survey.
- LaCharité, D. and P. Prévost (1999) The role of L1 and of teaching in the acquisition of English sounds by francophones, in A. Greenhill, H. Littlefield and C. Tano (eds) *Proceedings of BUCLD 23*. Somerville, MA: Cascadilla Press, pp. 373–385.
- Lacy, de P. (2006) *Markedness: Reduction and Preservation in Phonology*. Cambridge: Cambridge University Press.
- Lacy, de P. (ed.) (2007a) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press.
- Lacy, de P. (2007b) Themes in phonology, in P. de Lacy (eds) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, pp. 5–30.
- Ladd, D.R. (1996) *Intonational Phonology*. Cambridge: Cambridge University Press.
- Ladefoged, P. (1967) *Three areas of experimental phonetics*. London: Oxford University Press.
- Ladefoged, P. (1971) *Preliminaries to Linguistic Phonetics*. Chicago, IL: University of Chicago Press.
- Ladefoged, P. (1972) Phonetic prerequisites for a distinctive feature theory, in A. Valdman (ed.) *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*. The Hague: Mouton, pp. 273–286.
- Ladefoged, P. (1975) *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.
- Ladefoged, P. (1988) The many interfaces between phonetics and phonology. *UCLA Working Papers in Phonetics* 70: 13–23.
- Ladefoged, P. (1993) *A Course in Phonetics* (Third edition). New York: Harcourt, Brace, Jovanovich.

- Ladefoged, P. (1997) Instrumental techniques for linguistic fieldwork, in W.J. Hardcastle, J. William and J. Laver (eds) *The Handbook of Phonetic Sciences*. Oxford: Blackwell, pp. 137–166.
- Ladefoged, P. (2003) *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell.
- Ladefoged, P. and G. Loeb (2002) Preliminary studies on respiratory activity in speech. *UCLA Working Papers in Phonetics* 101: 50–60.
- Lado, R. (1957) *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor, MI: University of Michigan Press.
- Lahiri, A. and W.D. Marslen-Wilson (1991) The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition* 38: 245–294.
- Lahiri, A. and H. Reetz (2002) Underspecified recognition, in C. Gussenhoven, N. Werner and T. Rietveld (eds) *Labphon 7*. Berlin: Mouton, pp. 637–676.
- Lakoff, G. (1993) Cognitive phonology, in J. Goldsmith (ed.) *The Last Phonological Rule*. Chicago, IL: Chicago University Press, pp. 117–145.
- Laks, B. (2001) Un siècle de phonétique et de phonologie: quelques questions permanentes. *Modèles Linguistiques* 22: 75–103.
- Laks, B. (2006) La phonologie générative naturelle et la phonologie naturelle, in S. Aurooux, E.F.K. Koerner, H.-J. Niederehe and K. Versteegh (eds) *History of the Language Sciences* (Volume 3) (HSK 18.3). Berlin: Mouton de Gruyter, pp. 2271–2280.
- Lalonde, C.E. and J.F. Werker (1995) Cognitive influences on cross-language speech perception in infancy. *Infant Behaviour & Development* 18: 459–475.
- Langacker, R. (1987) *Foundations of Cognitive Grammar*, 2 Volumes. Stanford, CA: Stanford University Press.
- Larson, G. (1992) *Dynamic Computational Networks and the Representation of Phonological Information*. Ph.D. dissertation, University of Chicago.
- Laskaratou, C. (1989) *A Functional Approach to Constituent Order with Particular Reference to Modern Greek. Implications for Language Learning and Language Teaching*. Athens: Parousia Monograph Series 5.
- Laskaratou, C. (1998) Basic characteristics of Modern Greek word order, in A. Siewierska (ed.) *Constituent Order in the Languages of Europe*. Berlin: Mouton de Gruyter, pp. 151–174.
- Lass, R. (1984) *Phonology: An Introduction to Basic Concepts*. Cambridge: Cambridge University Press.
- Lavoie, L. (2001) *Consonant Strength: Phonological Patterns and Phonetic Manifestations*. New York: Garland Press.
- Leben, W. (1973) *Suprasegmental Phonology*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Leben, W. and O. Robinson (1977) Upside-down phonology. *Language* 53: 1–20.
- Lecours, A.R. and F. Lhermitte (1969) Phonemic paraphasias: Linguistic structures and tentative hypotheses. *Cortex* 5: 193–228.
- Lee, B.-G. (1982) A well-formedness condition on syllable structure, in I.S. Yang (ed.) *Linguistics in the Morning Calm*. Seoul: Hanshin Publishers, pp. 489–506.
- Legate, J.A. (2003) Some interface properties of the phase. *Linguistic Inquiry* 34: 506–516.
- Legendre, G., Y. Miyata and P. Smolensky (1990a) Can connectionism contribute to syntax? Harmonic Grammar, with an application, in M. Ziolkowski, M. Noske, and K. Deaton (eds) *Proceedings of the 26th Regional Meeting of the Chicago Linguistic Society*. Chicago, IL: Chicago Linguistic Society, pp. 237–252.
- Legendre, G., Y. Miyata and P. Smolensky (1990b) Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness: Theoretical

- foundations. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, pp. 388–395.
- Lehiste, I. (1970) *Suprasegmentals*. Cambridge, MA: MIT Press.
- Lenneberg, E.H. (1967) *Biological Foundations of Language*. Oxford: Wiley-Blackwell.
- Levelt, C.C. (1994) On the Acquisition of Place. Ph.D. dissertation, Leiden University, HIL dissertation series 8.
- Levelt, W.J.M. (1989) *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Levelt, W.J.M. and L. Wheeldon (1994) Do speakers have a mental syllabary? *Cognition* 50: 239–269.
- Levelt, W.J.M., A. Roelofs and A.S. Meyer (1999) A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22: 1–38.
- Levi, S. (2000) Modern Hebrew: A challenge to sympathy. Talk presented at the North American Phonology Conference 1, Montréal, April 2000.
- Levin, J. (1987) Between epenthetic and excrescent vowels, in *Proceedings of the Sixth West Coast Conference of Formal Linguistics*.
- Lew-Williams, C. and A. Fernald (2007) Young children learning Spanish made rapid use of the grammatical gender in spoken word recognition. *Psychological Science* 18: 193–198.
- Lieberman, A.M. and I.G. Mattingly (1985) The motor theory of speech perception revised. *Cognition* 21: 1–36.
- Lieberman, A.M., F.S. Cooper, D.P. Shankweiler and M. Studdert-Kennedy (1967) Perception of the speech code. *Psychological Review* 74: 431–461.
- Lieberman, A.M., K.S. Harris, J. Kinney and H. Lane (1961) The discrimination of relative onset-time of the components of certain speech and non-speech patterns. *Journal of Experimental Psychology* 61: 379–388.
- Lieberman, M. (1975) *The Intonational System of English*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Lieberman, M. (2002) Simple models for emergence of a shared vocabulary. Paper presented at *Laboratory Phonology VIII*, New Haven, CT.
- Lieberman, M. and A. Prince (1977) On stress and linguistic rhythm. *Linguistic Inquiry* 8: 249–336.
- Lichtheim, L. (1885) Über Aphasie. *Deutsches Archiv für Klinische Medizin* 36: 204–268.
- Liddell, S. (1980) *American Sign Language Syntax*. The Hague: Mouton.
- Liddell, S. (1984) THINK and BELIEVE: Sequentiality in American Sign Language. *Language* 60: 372–392.
- Liddell, S. (1990) Structures for representing handshape and local movement at the phonemic level, in S.D. Fischer and P. Siple (eds) *Theoretical Issues in Sign Language Research* (Volume 1). Chicago, IL: Chicago University Press, pp. 46–65.
- Liddell, S. and R. Johnson (1989) American Sign Language: The phonological base. *Sign Language Studies* 64: 197–277.
- Lightner, T. (1978) Generative phonology, in W.O. Dingwall (ed.) *A Survey of Linguistic Science* (Second edition). Stamford, CT: Greylock, pp. 1–32.
- Lightner, T. (1981) New explorations into derivational morphology, in D. Goyvaerts (ed.) *Phonology in the 80s*. Ghent: Story-Scientia, pp. 93–99.
- Lightner, T. (1985) On universal rules. *Linguisticae Investigationes* 9: 419–420.
- Liljencrants, J. and B. Lindblom (1972) Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48: 839–862.
- Lindblom, B. (1986) Phonetic universals in vowel systems, in J.J. Ohala and J.J. Jaeger (eds) *Experimental phonology*. Orlando, FL: Academic Press, pp. 13–44.
- Lindblom, B. (1990) Explaining phonetic variation: A sketch of the H and H theory, in W.J. Hardcastle and A. Marchal (eds) *Speech Production and Speech Modeling*. Dordrecht: Kluwer Academic Publishers, pp. 403–440.

- Lindblom, B., J. Lubker and T. Gay (1979) Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics* 7: 147–161.
- Lindblom, B., P. MacNeilage and M. Studdert-Kennedy (1984) Self-organizing processes and the explanation of language universals, in B. Butterworth, B. Comrie and O. Dahl (eds) *Explanations for Language Universals*. Berlin: Walter de Gruyter and Co., pp. 181–203.
- Lisker, L. (1975) Is it VOT or a first-formant transition detector? *Journal of Acoustical Society of America* 57: 1547–1551.
- Lisker, L. and A.S. Abramson (1964) A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20: 384–422.
- Lisker, L. and A.S. Abramson (1970) The voicing dimension: Some experiments in comparative phonetics, *Proceedings of the 6th International Congress of Phonetic Sciences*. Prague: Academia Publishing House, pp. 563–567.
- Locke, J.L. (1993) *The Child's Path to Spoken Language*. Cambridge: Harvard University Press.
- Lombardi, L. (1991a) Laryngeal Features and Laryngeal Neutralization. Ph.D. dissertation, University of Massachusetts, Amherst.
- Lombardi, L. (1991b) Laryngeal Features and Privativity. Ms., University of Massachusetts, Amherst.
- Lombardi, L. (1997) Coronal epenthesis and markedness. Ms., University of Maryland. [ROA#245].
- Lombardi, L. (2001a) Introduction. Segmental phonology in Optimality Theory, in L. Lombardi (ed.) *Constraints and Representations*. Cambridge: Cambridge University Press, pp. 1–9.
- Lombardi, L. (2001b) Why place and voice are different. Constraint-specific alternations in Optimality Theory, in L. Lombardi (ed.) *Segmental Phonology in Optimality Theory: Constraints and Representations*. Cambridge: Cambridge University Press, pp. 13–45.
- Lowenstamm, J. (1981) On the maximal cluster approach to syllable structure. *Linguistic Inquiry* 12: 575–604.
- Lowenstamm, J. (1996) CV as the only syllable type. Current trends in phonology, in J. Durand and B. Laks (eds) *Models and Methods*. Salford, Manchester: ESRI, pp. 419–441.
- Lowenstamm, J. (2006) Développement de la phonologie métrique, in S. Auroux, E.F.K. Koerner, H-J Niederehe and K. Versteegh (eds) *History of the Language Sciences* (Volume 3) (HSK 18.3). Berlin: Mouton de Gruyter, pp. 2303–2312.
- Lowenstamm, J. and J.D. Kaye (1986) Compensatory lengthening in Tiberian Hebrew, in L. Wetzels and E. Sezer (eds) *Studies in Compensatory Lengthening*. Dordrecht: Foris Publications, pp. 97–132.
- Łubowicz, A. (2003) Contrast Preservation in Phonological Mappings. Ph.D. dissertation, University of Massachusetts, Amherst. Rutgers Optimality Archive #554. Available at www.roa.rutgers.edu.
- Łubowicz, A. (2005) Locality of conjunction, in J. Alderete, C. hye Han and A. Kochetov (eds) *Proceedings of the 24th West Coast Conference on Formal Linguistics*. Cambridge, MA: Cascadilla Press, pp. 254–262.
- Luce, P.A. and D.B. Pisoni (1998) Recognizing spoken words: the neighborhood activation model. *Ear and Hearing* 19: 1–36.
- MacWhinney, B. (2000) *The CHILDES Project: Tools for Analyzing Talk* (Third edition). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mah, J. and J. Archibald (2003) Acquisition of L2 length contrasts, in J.M. Liceras, H. Zobl and H. Goodluck (eds) *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*. Somerville, MA: Cascadilla, pp. 208–212.

Bibliography

- Major, R.C. (2001) *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Major, R.C. and E. Kim (1999) The similarity differential rate hypothesis, in J. Leather (ed.) *Phonological Issues in Language Learning: A Supplement to Language Learning* 49, Supplement 1. Ann Arbor, MI: Basil Blackwell, pp. 151–183.
- Malikouti-Drachman, A. and G. Drachman (1989) *Tonismos sta Ellinika [Stress in Greek]*, in *Studies in Greek Linguistics*. University of Thessaloniki, pp. 127–143.
- Mandel, M. (1981) *Phonotactics and Morphophonology in American Sign Language*. Ph.D. dissertation, University of California.
- Mani, N. and K. Plunkett (2007) Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language* 57: 252–272.
- Mani, N. and K. Plunkett (2008a) 14-month-olds pay attention to vowels in novel words. *Developmental Science* 11: 53–59.
- Mani, N. and K. Plunkett (2008b) Graded Sensitivity to Mispronunciations of Vocalic Features in Early Words *Proceedings of the 32nd Annual Boston University Conference on Language Development*. Boston, MA: Cascadilla Press.
- Mani, N. and K. Plunkett (2010) 12-month-olds know their cups from their keps and tups. *Infancy*, 15 (5): 445–470.
- Mani, N., J. Coleman and K. Plunkett (2008) Phonological specificity of vocalic features at 18-months. *Language and Speech* 51: 3–21.
- Manuel, S. (1990) The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America* 88: 1286–1298.
- Manuel, S. (1999) Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts, in W.J. Hardcastle and N. Hewlett (eds) *Coarticulation: Theory, Data and Techniques*. Cambridge: Cambridge University Press, pp. 179–198.
- Manuel, S. and R. Krakow (1984) Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status reports on Speech Research*, SR-77/78: 69–78.
- Marantz, A. (1988) Clitics, morphological merger and the mapping to phonological structure, in M. Hammond and M. Noonan (eds) *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego, CA: Academic Press, pp. 253–270.
- Marcus, G., S. Pinker, M. Ullman, M. Hollander, J. Rosen and F. Xu (1992) Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 57 (4).
- Marotta, G. (2008) Lenition in Tuscan Italian (Gorgia Toscana), in Carvalho et al. (eds) *Lenition and Fortition*. Berlin and New York: Mouton de Gruyter, pp. 235–271.
- Marslen-Wilson, W. and A. Welsh (1978) Processing interactions during word-recognition in continuous speech. *Cognitive Psychology* 10: 29–63.
- Martinet, A. (1952) Function, structure, and sound change. *Word* 8: 1–32.
- Mascaró, J. (1976) Catalan Phonology and the Phonological Cycle. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Massaro, D.W. and M.M. Cohen (1983) Phonological constraints in speech perception. *Perception and Psychophysics* 34: 338–348.
- Matthews, J. (1997) The influence of pronunciation training on the perception of second language contrasts, in J. Leather and A. James (eds) *New Sounds* 97. Klagenfurt: University of Klagenfurt Press, pp. 223–239.
- Mattingly, I.G. and A.M. Liberman (1969) The speech code and the physiology of language, in K.N. Leibovic (ed.) *Information Processing in the Nervous System*. New York: Springer.
- Mattys, S.L. and P.W. Jusczyk (2001) Phonotactic cues for segmentation of fluent speech by infants. *Cognition* 78: 91–121.

- Maye, J., D.J. Weiss and R.N. Aslin (2008) Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science* 11: 122–134.
- Maye, J., J.F. Werker and L. Gerken (2002) Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82: B101–B111.
- McCarthy, J.J. (1979a) Formal Problems in Semitic Phonology and Morphology. Ph.D. dissertation, Massachusetts Institute of Technology. Published by Garland Press, New York, 1985.
- McCarthy, J.J. (1979b) On stress and syllabification. *Linguistic Inquiry* 10: 443–466.
- McCarthy, J.J. (1986) OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17: 207–263.
- McCarthy, J.J. (1988) Feature geometry and dependency: A review. *Phonetica* 45: 84–108.
- McCarthy, J.J. (1993) A case of surface constraint violation. *Canadian Journal of Linguistics* 38 (2): 169–195.
- McCarthy, J.J. (1996) Remarks on phonological opacity in Optimality Theory, in J. Lecarme, J. Lowenstamm, and U. Shlonsky (eds) *Studies in Afroasiatic Grammar: Papers from the 2nd Conference on Afroasiatic Linguistics, Sophia Antipolis, 1994*. The Hague: Holland Academic Graphics, pp. 215–243.
- McCarthy, J.J. (1998) Sympathy and phonological opacity. Ms., University of Massachusetts, Amherst.
- McCarthy, J.J. (1999) Sympathy and phonological opacity. *Phonology* 16: 331–399.
- McCarthy, J.J. (2002) *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.
- McCarthy, J.J. (2007) *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.
- McCarthy, J.J. (2007a) Derivations and levels of representation, in P. de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, pp. 99–117.
- McCarthy, J.J. (2007b) Restraint of analysis, in M. Krämer, S. Blaho, and P. Bye (eds) *Freedom of Analysis?* Berlin: Mouton de Gruyter, pp. 203–232.
- McCarthy, J.J. (2008a) The gradual path to cluster simplification. *Phonology* 25: 271–319.
- McCarthy, J.J. (2008b) The serial interaction of stress and syncope. *Natural language and Linguistic Theory* 26: 499–546.
- McCarthy, J.J. (2009) Harmony in harmonic serialism. Ms., University of Massachusetts, Amherst [ROA#1009].
- McCarthy, J.J. and A. Prince (1986) Prosodic Morphology. Ms., Rutgers University.
- McCarthy, J.J. and A. Prince (1993a) Generalized alignment, in G. Booij and J. van Marle (eds) *Yearbook of Morphology 1993*. Dordrecht: Kluwer Academic Publishers, pp. 79–153.
- McCarthy, J.J. and A. Prince (1993b) Prosodic morphology I: Constraint interaction and satisfaction. Ms., University of Massachusetts and Rutgers University.
- McCarthy, J.J. and A. Prince (1994) The emergence of the unmarked: Optimality in prosodic morphology. Ms., University of Massachusetts, Amherst and Rutgers University.
- McCarthy, J.J. and A. Prince (1995) Faithfulness and reduplicative identity, in J. Beckman, L. Dickey and S. Urbanczyk (eds) *Papers in Optimality Theory, University of Massachusetts Occasional Papers* 18: 249–384.
- McCarthy, J.J. and A. Prince (1999) Faithfulness identity in Prosodic Morphology, in R. Kager, H. van der Hulst and W. Zonnefeld (eds) *The Prosody-Morphology Interface*. Cambridge: Cambridge University Press, pp. 218–309.
- McCarthy, J.J. and A. Prince (2001) Prosodic Morphology. Constraint interaction and satisfaction. Ms., [ROA#482], Rutgers University.
- McCawley, J.D. (1965) *The Accentual System of Standard Japanese*. Ph.D. thesis, Massachusetts Institute of Technology.
- McCawley, J. (1968) *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton de Gruyter.

- McClelland, J. and D.D.E. Rumelhart (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review* 88: 375–407.
- McClelland, J. and J.L. Elman (1986) The TRACE model of speech perception. *Cognitive Psychology* 18: 1–86.
- McKeever Dauer, Rebecca (1980) Stress and Rhythm in Modern Greek. Ph.D. dissertation, University of Edinburgh.
- McMahon, A. (2000) *Lexical Phonology and the History of English*. Cambridge: Cambridge University Press.
- McMahon, A. (2007) Who's afraid of the vowel shift rule? *Language Sciences* 29: 341–359.
- McMurray, B. and R.N. Aslin (2004) Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy* 6: 203–229.
- McMurray, B. and R.N. Aslin (2005) Infants are sensitive to within-category variation in speech perception. *Cognition* 95: B15–B26.
- McMurray, B., M. Spivey and R.N. Aslin (2000) The perception of consonants by adults and infants: Categorical or categorized? Preliminary results, in K. Crosswhite and J. Magnuson (eds) *The University of Rochester Working Papers in the Language Sciences*. pp. 215–256.
- McQueen, J. and A. Cutler (1997) Cognitive processes in speech perception, in W.J. Hardcastle and J. Laver (eds) *The handbook of phonetic sciences*. Oxford: Blackwell, pp. 566–586.
- Medin, D.L. (1983) Structural principles of categorization, in T. Tighe and B. Shepp (eds) *Perception, Cognition, and Development: Interactional Analyses*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 203–230.
- Medin, D.L. and M.M. Schaffer (1978) Context theory of classification learning. *Psychological Review* 85: 207–238.
- Meek, J. (2002) Basic principles of optical imaging and application to the study of infant development. *Developmental Science* 5: 371–380.
- Mehler, J. (1985) Language related dispositions in early infancy, in J. Mehler and R. Fox (eds) *Neonate Cognition: Beyond the Blooming, Buzzing Confusion*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 7–28.
- Mehler, J. and A. Christophe (1995) Maturation & learning of language in the first year of life, in M.S. Gazzaniga (ed.) *The Cognitive Neuro-Sciences: A Handbook for the Field*. Cambridge, MA: MIT Press, pp. 943–954.
- Mehler, J., P. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini and C. Amiel-Tison (1988) A precursor of language acquisition in young infants. *Cognition* 29: 143–178.
- Mehler, J., M. Nespore, J. Gervain, A.D. Endress and Shukla M. (2008) Mechanisms of language acquisition: Imaging and behavioral evidence, in C.A. Nelson and M. Luciana (eds) *Handbook of Developmental Cognitive Neuroscience* (Second edition). Cambridge, MA: MIT Press, pp. 325–336.
- Meier, R., K. Cormier and D. Quinto-Pozos (eds) (2002) *Modality in Language and Linguistic Theory*. Cambridge: Cambridge University Press.
- Melvold, J.L. (1990) Structure and stress in the phonology of Russian. PhD dissertation, Massachusetts Institute of Technology.
- Mester, R.A. (1994) The quantitative trochee in Latin, *Natural Language and Linguistic Theory* 12: 1–61.
- Metsala, J.L. (1999) Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology* 91: 3–19.
- Mielke, J. (2004) The Emergence of Distinctive Features. Ph.D. dissertation, Ohio State University.
- Mielke, J., M. Armstrong, and E. Hume (2003) Looking through opacity. *Theoretical Linguistics* 29: 123–139.

- Milroy, J. and J. Harris (1980) When is a merger not a merger? The MEAT/MATE problem in a present-day English vernacular. *English World-Wide* 1: 199–210.
- Mohannan, K. (1986) *The Theory of Lexical Phonology*. Dordrecht: Reidel.
- Mohanan, K. (2000) The theoretical substance of the optimality formalism. *The Linguistic Review* 17: 143–166.
- Monachesi, P. (1999) Syntactic and prosodic properties of Italian restructuring verbs, in P. Tamanji, M. Hirotani and N. Hall (eds) *Proceedings of NELS 29*. Amherst: GLSA, pp. 277–291.
- Mooshammer, C. and S. Fuchs (2002) Stress distinction in German: Simulating kinematic parameters of tongue tip gestures. *Journal of Phonetics* 30: 337–355.
- Moreton, E. (2002) Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84: 55–71.
- Moreton, E. (2008) Analytic bias and phonological typology. *Phonology* 25: 83–127.
- Moreton, E. and P. Smolensky (2002) Typological consequences of local constraint conjunction. Ms., John Hopkins University.
- Mowrey, R.A. and I.R. MacKay (1990) Phonological primitives: Electromyographic speech error evidence. *Journal of the Acoustical Society of America* 88: 1299–1312.
- Muñoz, C. (2008) Age-related differences in foreign language learning. Revisiting the empirical evidence. *International Review of Applied Linguistics* 46: 197–220.
- Munro, M.J. (1998) The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition* 20: 139–154.
- Munro, M.J. and T.M. Derwing (1999) Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 49: 285–310.
- Munro, M.J., T.M. Derwing and S.L. Morton (2006) The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition* 28: 111–131.
- Munson, B., B.A. Kurtz and J. Windsor (2005) The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research* 48: 1033–1047.
- Myers, S. (1987) Vowel shortening in English. *Natural Language and Linguistic Theory* 5: 485–518.
- Myers, S. (1991), Persistent rules. *Linguistic Inquiry* 22: 315–344.
- Nam, H. (2007) A competitive, coupled oscillator model of moraic structure: Split-gesture dynamics focusing on positional asymmetry, in J. Cole and J. Hualde (eds) *Laboratory Phonology IX*. Berlin and New York: Mouton de Gruyter, pp. 483–506.
- Napoli, D.J. and M. Nespore (1979) The syntax of word-initial consonant gemination in Italian. *Language* 55: 812–841.
- Narayan, C.R., J.F. Werker and P.S. Beddor (2010) The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, 13 (3), 407–420.
- Nasukawa, K. (1995) Melodic structure and no constraint-ranking in Japanese verbal inflexion. Paper presented at the Autumn Meeting of the Linguistic Association of Great Britain, University of Essex.
- Nasukawa, K. (1998) An integrated approach to nasality and voicing, in E. Cyran (ed.) *Structure and Interpretation: Studies in Phonology*. Lublin: Folium, pp. 205–225.
- Nasukawa, K. (2005) *A Unified Approach to Nasality and Voicing*. Berlin and New York: Mouton de Gruyter.
- Nasukawa, K. and P. Backley (2005) Dependency relations in Element Theory. *Leiden Papers in Linguistics* 2 (4): 77–93.
- Nasukawa, K. and P. Backley (2008) Affrication as a performance device. *Phonological Studies* 11: 35–46.

- Nation, K., C.M. Marshall and G.T. Altmann (2003) Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology* 86: 314–329.
- Nazzi, T. (2005) Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition* 98: 13–30.
- Nazzi, T., J. Bertoncini and J. Mehler (1998) Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception & Performance* 24: 756–766.
- Neeleman, A. and H. van de Koot (2006) On syntactic and phonological representations. *Lingua* 116: 1524–1552.
- Nespor, M. (1999) The phonology of clitic groups, in H. van Riemsdijk (ed.) *Clitics in the Languages of Europe*. Berlin/New York: Mouton de Gruyter, pp. 865–887.
- Nespor, M. and I. Vogel (1982) Prosodic domains of external sandhi rules, in H. van der Hulst and N. Smith (eds) *The Structure of Phonological Representations*. Part I. Dordrecht: Foris Publications, pp. 225–255.
- Nespor, M. and I. Vogel (1983) Prosodic structure above the word, in A. Cutler and R. Ladd (eds) *Prosody: Models and Measurements*. Berlin: Springer, pp. 123–140.
- Nespor, M. and I. Vogel (1986) *Prosodic Phonology*. Dordrecht: Foris Publications.
- Nespor, M., M. Pena and J. Mehler (2003) On the different role of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio* 2: 221–247.
- Nespoulous, J.-L., R. Béland, D. Caplan and A.R. Lecours (1984) Phonological disturbances in aphasia: Is there a 'markedness effect' in aphasic phonemic errors?, in F.C. Rose (ed.) *Advances in Aphasiology (Vol. 42): Progress in Aphasiology*. London: Raven Press, 203–214.
- Nespoulous, J.-L., Y. Joannette, B. Ska, D. Caplan and A.R. Lecours (1987) Production deficits in Broca's and conduction aphasia: Repetition vs. reading, in E. Keller and M. Gopnik (eds) *Motor and Sensory Processes in Language*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Newkirk, D. (1998) On the temporal organisation of movement in American Sign Language. *Sign Language and Linguistics* 1: 173–211.
- Newkirk, D., E. Klima, C. Pedersen and U. Bellugi (1980) Linguistic evidence from slips of the hand, in V.A. Fromkin (ed.) *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand*. New York: Academic Press, pp. 165–198.
- Newman, S. (1944) *The Yokuts Language of California*, [=The Viking Fund Publications in Anthropology, 2]. The Viking Fund.
- Newmeyer, F. (1986) *Linguistic Theory in America* (Second edition). New York: Academic Press.
- Newmeyer, F. (1996) *Generative Linguistics: A Historical Perspective*. London: Routledge.
- Newmeyer, F. (1998) *Language Form and Language Function*. Cambridge, MA: MIT Press.
- Newmeyer, F. (2005) Some remarks on Roeper's remarks on Chomsky's 'Remarks'. *SKASE Journal of Theoretical Linguistics* 2: 26–39. [www.skase.sk].
- Newton, B. (1972) *The Generative Interpretation of Dialect. A Study of Modern Greek Phonology*. Cambridge: Cambridge University Press.
- Nikiema, E. (1989a) Gouvernement propre et licenciement en phonologie: le cas du tangale. *Langues orientales anciennes philologie et linguistique* 2: 225–251.
- Nikiema, E. (1989b) Vocalic epenthesis reanalyzed: The case of Tangale, in J. Hutchinson and V. Manfredi (eds) *Current Approaches to African Linguistics*, 7. Dordrecht: Foris, pp. 41–51.
- Nordhoff, S. (2009) *A Grammar of Upcountry Sri Lanka Malay*. Ph.D. dissertation, LOT Dissertation series No. 226. Utrecht: LOT.
- Norris, D. (1994) Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52: 189–234.

- Norris, D., J. McQueen, A. Cutler and S. Butterfield (1997) The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology* 34: 191–243.
- Nosofsky, R.M. (1986) Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115: 39–57.
- Nosofsky, R.M. (1988) Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14: 700–708.
- Nowak, M.A. and D. Krakauer (1999) The evolution of language. *Proceedings of the National Academy of Sciences* 96: 8028–8033.
- Nunes, J. (2004) *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.
- Nyst, Victoria (2007) A Descriptive Analysis of Adamorobe Sign Language. Ph.D. dissertation, University of Amsterdam.
- Odden, D. (1987) Kimatuumbi phrasal phonology. *Phonology* 4: 13–26.
- Odden, D. (1988) Anti antigemination and the OCP. *Linguistic Inquiry* 19: 451–475.
- Odden, D. (1990) Syntax, lexical rules and postlexical rules in Kimatuumbi, in S. Inkelas and D. Zec (eds) *The Phonology-Syntax Connection*. Chicago, IL: University of Chicago Press, pp. 259–277.
- Odden, D. (2008) Ordering, in B. Vaux and A. Nevins (eds) *Rules, Constraints, and Phonological Phenomena*. Oxford: Oxford University Press, pp. 61–120.
- Ogden, R. (1993) What Firthian Prosodic Analysis has to say to us, in T.M. Ellison and J.M. Scobbie (eds) *Computational Phonology (Edinburgh Working Papers in Cognitive Science)* 8. Edinburgh: Edinburgh University, pp. 107–127.
- Oh, J., S.-A. Jun, L. Knightly and T.K. Au (2003) Holding on to childhood language memory. *Cognition* 86: B53–B64.
- Ohala, D. (1999) The influence of sonority on children's clusters reductions. *Journal of Communication Disorders* 32: 397–422.
- Ohala, D. (2008) Phonological acquisition in a first language, in M. Zampini and J. Hansen (eds) *Phonology and Second Language Acquisition*. Amsterdam: John Benjamins, pp. 19–29.
- Ohala, J.J. (1981) The listener as a source of sound change, in C.S. Masek, R.A. Hendrick and M.F. Miller (eds) *Proceedings of the Chicago Linguistics Society 17: Papers from the Parasession on Language and Behavior*. Chicago, IL: University of Chicago Press, pp. 178–203.
- Ohala, J.J. (1983) The origin of sound patterns in vocal tract constraints, in P.F. MacNeilage (ed.) *The Production of Speech*. New York: Springer, pp. 189–216.
- Ohala, J.J. (1989) Sound change is drawn from a pool of synchronic variation, in L.E. Breivik and E.H. Jahr (eds) *Language Change: Contributions to the Study of its Causes*. Berlin: Mouton de Gruyter, pp. 173–198.
- Ohala, J.J. (1992) Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Proceedings of the Chicago Linguistic Society* 26: 319–338.
- Ohala, J.J. (1993) The phonetics of sound change, in C. Jones (ed.) *Historical Linguistics: Problems and Perspectives*. London: Longman, pp. 237–278.
- Ohala, J.J. (2005) Phonetic explanations for sound patterns: Implications for grammars of competence, in W.J. Hardcastle and M. Beck (eds) *A Figure of Speech: A Festschrift for John Laver*. London: Lawrence Erlbaum Associates, pp. 23–38.
- Öhman, S.E.G. (1966) Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America* 39: 151–168.
- Oller, D.K. (1980) The emergence of the sounds of speech in infancy, in G. Yeni-Komshian, J. Kavanagh and C. Ferguson (eds) *Child phonology: Production 1*. New York: Academic Press, pp. 93–112.

- Oller, D.K., R.E. Eilers, A.R. Neal and H.K. Schwartz (1999) Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders* 32: 223–245.
- Oller, D.K., R.E. Eilers, R. Urbano and A.B. Cobo-Lewis (1997) Development of precursors to speech in infants exposed to two languages. *Journal of Child Language* 24: 407–425.
- Olney, R.L. and E.K. Scholnik (1976) Adult judgments of age and linguistic differences in infant vocalization. *Journal of Child Language* 3: 145–155.
- Olson, K. and J. Hajek (2004) A crosslinguistic lexicon of the labial flap. *Linguistic Discovery* 2 (2): 21–57. [www.linguistic-discovery.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/2/xmlpage/1/article/262].
- Onaka, A. (2003) Domain-initial strengthening in Japanese: An acoustic and articulatory study. *Proceedings of the 15th International Congress of Phonetic Sciences*, pp. 2091–2094.
- Onaka, A., C. Watson, S. Palethorpe and J. Harrington (2003) An acoustic analysis of domain-initial strengthening effect in Japanese, in S. Palethorpe and M. Tabain (eds) *Proceedings of the 6th International Seminar on Speech Production*, Sydney: [. . .], pp. 201–206.
- Oostendorp, M. van (2002) The phonological and morphological status of the prosodic word adjunct. *Linguistische Berichte, Sonderheft* 11: 209–235.
- Oostendorp, M. van (2003) Ambisyllabicity and fricative voicing in West-Germanic dialects, in C. Féry and R. van de Vijver (eds) *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press, pp. 304–337.
- Oostendorp, M. van (2005) The first person singular in Dutch dialects, in L. Bateman and C. Ussery (ed.) *Proceedings of the Thirty-Fifth Annual Meeting of the North East Linguistic Society*. Amherst, MA: GLSA, pp. 1–12.
- Oostendorp, M. van (2006) A theory of morphosyntactic colours. Ms., Meertens Instituut, Amsterdam.
- Oostendorp, M. van (2008) Incomplete devoicing in formal phonology. *Lingua* 118: 1362–1374.
- Oostendorp, M. van and J. van de Weijer (2005) Phonological alphabets and the structure of the segment, in M. van Oostendorp and J. van de Weijer (eds) *The Internal Organisation of Phonological Segments*. Berlin: Mouton de Gruyter, pp. 1–23.
- Orfanidou, E., R. Adam, G. Morgan and J. McQueen (2010) Recognition of signed and spoken language: Different sensory inputs, the same segmentation procedure. *Journal of Memory and Language* 62: 272–283.
- Padgett J. (1995) *Stricture in Feature Geometry*. Stanford, CA: CSLI.
- Padgett, J. (2001a) Constraint conjunction vs. grounded constraint subhierarchies. Ms., University of California, Santa Cruz.
- Padgett, J. (2001b) Contrast dispersion and Russian palatalization, in E. Hume and K. Johnson (eds) *The Role of Speech Perception in Phonology*. San Diego, CA: Academic Press, pp. 187–218.
- Padgett, J. (2003a) Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory* 21: 39–87.
- Padgett, J. (2003b) The emergence of contrastive palatalization in Russian, in E. Holt (ed.) *Optimality Theory and Language Change*. Dordrecht: Kluwer Academic Press, pp. 307–335.
- Pallier, C., A. Colomé and N. Sebastián-Gallés (2001) The influence of native-language phonology on lexical access: Exemplar-based vs. abstract lexical entries. *Psychological Science* 12: 445–449.
- Paradis, C. (1988) On constraints and repair strategies. *The Linguistic Review* 6: 71–97.

- Paradis, C. (1996) The inadequacy of filters and faithfulness in loanword adaptation, in J. Durand and B. Laks (eds) *Current Trends in Phonology: Models and Methods*. Salford: University of Salford Publications. ESRI, pp. 509–534.
- Paradis, C. and F. El Fenne (1995) French verbal inflection revisited: Constraints, repairs and floating consonants. *Lingua* 95: 169–204.
- Parisse, C. and E. Soubeyrand (2003) Computerized test for oral language perception: TIPLo. *Journal of Multilingual Communication Disorders* 1: 53–61.
- Parker, S. (2009) The transparency and density of phonological rule interactions. *Linguistic Typology* 13: 197–265.
- Passy, P. (1891) *Etude sur les changements phonétiques et leurs caractères généraux*. Paris: Librairie Firmin-Didot.
- Pate, D.S., E.M. Saffran and N. Martin (1987) Specifying the nature of the production impairment in a conduction aphasic: A case study. *Language and Cognitive Processes* 2: 43–84.
- Pater, J. (1993) Theory and methodology in the study of metrical parameter (re)setting. *McGill Working Papers in Linguistics / Cahiers Linguistiques de McGill* 9: 211–243.
- Pater, J. (1997) Metrical parameter missetting in second language acquisition, in S.J. Hannahs and M. Young-Scholten (eds) *Focus on Phonological Acquisition*. Amsterdam: John Benjamins, pp. 235–261.
- Pater, J. (2003) The perceptual acquisition of Thai phonology by English speakers: Task and stimulus effects. *Second Language Research* 19: 209–223.
- Pater, J. (2004) Bridging the gap between perception and production with minimally violable constraints, in R. Kager, J. Pater and W. Zonneveld (eds) *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, pp. 219–244.
- Pater, J. (2009) Weighted constraints in generative linguistics. *Cognitive Science* 33: 999–1035.
- Paul, H. (1880) *Prinzipien der Sprachgeschichte*. Halle: Niemeyer. [Translated 1970 by H.A. Strong as *Principles of the History of Language*. College Park: McGrath].
- Pearce, M. (2007) The Interaction of Tone with Voicing and Foot Structure: Evidence from Kera Phonetics and Phonology. Ph.D. dissertation, University College London.
- Peña, M., A. Maki, D. Kovacic, G. Dehaene-Lambertz, H. Koizumi, F. Bouquet and J. Mehler (2003) Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of the United States of America* 100: 11702–11705.
- Peperkamp, S. (1997) Prosodic Words. Ph.D. dissertation, University of Amsterdam.
- Peperkamp, S. (2004) Lexical exceptions in stress systems: Arguments from early language acquisition and adult speech perception. *Language* 80: 98–126.
- Peperkamp, S. and E. Dupoux (2002) A typological study of stress ‘deafness’, in C. Gussenhoven and N. Warner (eds) *Papers in Laboratory Phonology 7*. Berlin: Mouton de Gruyter, pp. 203–240.
- Peperkamp, S., E. Dupoux and N. Sebastián-Galles (1999) Perception of stress by French, Spanish, and bilingual subjects. *Proceedings of EuroSpeech '99* 6: 2683–2686.
- Perlmutter, D. (1992) Sonority and syllable structure in American Sign Language. *Linguistic Inquiry* 23: 407–442.
- Pesetsky, D. (1996) Optimality Theory and syntax: Movement and pronunciation, in D. Archangeli and T. Langedoen (eds) *Optimality Theory*. Oxford: Blackwell, pp. 134–170.
- Pesetsky, D. (1998) Some optimality principles of sentence pronunciation, in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky (eds) *Is the best good enough? Optimality and Competition in Syntax*. Cambridge, MA: MIT Press & MIT Working Papers in Linguistics, pp. 337–383.

- Philippaki-Warbuton, I. (1970) *On the Verb in Modern Greek*. Bloomington, IN: Indiana University Publications.
- Philippaki-Warbuton, I. (1976) On the boundaries of morphology and phonology: A case study from Modern Greek. *Journal of Linguistics* 12: 259–278.
- Philippaki-Warbuton, I. (1985) Word order in Modern Greek. *Transactions of the Philological Society*. Salford: ESRI, 113–143.
- Philippaki-Warbuton, I. (1987) The theory of empty categories and the pro-drop parameter in Modern Greek. *Journal of Linguistics* 23: 289–318.
- Philippaki-Warbuton, I. (1998) Functional categories and Modern Greek syntax. *The Linguistic Review* 15: 58–186.
- Philippaki-Warbuton I. and V. Spyropoulos (1999) On the boundaries of inflection and syntax, in G. Booij and J. van Marle (eds) *The Yearbook of Morphology 1998*. Dordrecht: Kluwer Academic Press, pp. 45–72.
- Phillips, C. (1996) Order and Structure. Ph.D. dissertation, Massachusetts Institute of Technology.
- Piaget, J. (1954) *Intelligence and Affectivity: Their Relationship during Child Development*. Palo Alto, CA: Annual Review, Inc.
- Pierrehumbert, J.B. (1980) The Phonology and Phonetics of English Intonation. Ph.D. dissertation, Massachusetts Institute of Technology.
- Pierrehumbert, J.B. (1990) Phonological and phonetic representation. *Journal of Phonetics* 28: 375–394.
- Pierrehumbert, J.B. (2001) Exemplar dynamics: Word frequency, lenition, and contrast, in J. Bybee and P.J. Hopper (eds) *Frequency Effects and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins, pp. 137–157.
- Pierrehumbert, J.B. (2002) Word-specific phonetics, in C. Gussenhoven and N. Warner (eds) *Laboratory phonology VII*. Berlin: Mouton de Gruyter, pp. 101–140.
- Pierrehumbert, J.B. (2003) Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46: 115–154.
- Pierrehumbert, J.B. and M.E. Beckman (1988) *Japanese Tone Structure*. Cambridge, MA: MIT Press.
- Pierrehumbert, J.B. and J. Hirschberg (1990) The meaning of intonational contours in the interpretation of discourse, in R.R. Cohen, J. Morgan and M.E. Pollack (eds) *Intentions in Communication*. Cambridge, MA: MIT Press, pp. 271–311.
- Pierrehumbert, J.B. and D. Talkin (1992) Lenition of /h/ and glottal stop, in G. Docherty and D.R. Ladd (eds) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press, pp. 90–117.
- Piggott, G.L. (1992) Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory* 10: 33–72.
- Piggott, G.L. (1997) Licensing and alignment: A conspiracy in harmony. *Phonology* 14: 437–477.
- Piggott, G.L. (1999) At the right edge of words. *The Linguistic Review* 16: 143–185.
- Piggott, G. and H. van der Hulst (1997) Locality and the nature of nasal harmony. *Lingua* 103: 85–112.
- Pike, K. and E. Pike (1947) Immediate constituents of Mazatec syllables. *International Journal of American Linguistics* 13: 78–91.
- Pike, K., B. Ralph and B. Bascum (1959) Instrumental Collaboration on a Tepehuan (Uto Aztecan) Pitch Problem. *Phonetica* 3: 1–22.
- Pisoni, D.B., R.N. Aslin, A.J. Pery and B.L. Hennessy (1982) Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* 8: 297–314.

- Pitt, M.A. (1998) Phonological processes and the perception of phonetically illegal consonant clusters. *Perception and Psychophysics* 60: 941–951.
- Pitt, M.A. and J.M. McQueen (1998) Is compensation for coarticulation mediated between monosyllabic spoken words? *Journal of Experimental Psychology: Learning, Memory and Cognition* 15: 378–387.
- Plaut, D.C., J. McClelland, M.S. Seidenberg and K. Patterson (1996) Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review* 103: 56–115.
- Ploch, S. (1999) Nasals on My Mind: The Phonetic and the Cognitive Approach to the Phonology of Nasality. Ph.D. dissertation, SOAS.
- Plug, L. (2005) From words to actions: The phonetics of *eigenlijk* in two pragmatic contexts. *Phonetica* 62: 131–145.
- Plug, L. (2010) Pragmatic constraints in usage-based phonology, with reference to some Dutch phrases. *Journal of Pragmatics* 42 (7): 2014–2035.
- Pöchtrager, M. (2006) The Structure of Length. Ph.D. dissertation, University of Vienna.
- Poeck, K. (1983) What do we mean by ‘aphasic syndromes’? A neurologist’s view. *Brain and Language* 20: 79–89.
- Polgárdi, K. (1999) Constraint ranking, government licensing and the fate of final empty nuclei, in J. Rennison and K. Kühnhammer (eds) *Phonologica 1996*. The Hague: Holland Academic Graphics, pp. 167–182.
- Poliquin, G. (2006) Canadian French Vowel Harmony. Ph.D. dissertation, Harvard University.
- Polka, L. (1991) Cross-language speech perception in adults: Phonemic, phonetic and acoustic contributions. *Journal of the Acoustical Society of America* 89: 2961–2977.
- Pollard, C. and I. Sag (1994) *Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI.
- Poplack, S. (1980a) The notion of the plural in Puerto Rican Spanish: Competing constraints on (s) deletion, in W. Labov (ed.) *Locating Language in Time and Space*. New York: Academic Press, pp. 55–67.
- Poplack, S. (1980b) Deletion and disambiguation in Puerto Rican Spanish. *Language* 56: 371–385.
- Port, R. and M.L. O’Dell (1985) Neutralization of syllable-final voicing in German. *Journal of Phonetics* 13: 455–471.
- Port, R. and P. Crawford (1989) Incomplete neutralization and pragmatics in German. *Journal of Phonetics* 17: 257–282.
- Poser, W.J. (1990) Evidence for foot structure in Japanese. *Language* 66 (1): 78–105.
- Poser, W. (1986) Diyari stress, metrical structure assignment and the nature of metrical representation. *Proceedings of the West Coast Conference on Formal Linguistics* 5: 178–191.
- Postal, P. (1968) *Aspects of Phonological Theory*. New York: Harper and Row.
- Prieto, P. (1997) Prosodic manifestation of syntactic structure in Catalan, in F. Martínez-Gil and A. Morales-Front (eds) *Issues in the Phonology of the Major Iberian Languages*. Washington, DC: Georgetown University Press, pp. 179–199.
- Prince, A. (1985) *Improving Tree Theory*. *Proceedings of the 11th Annual Meeting of the Berkeley Linguistics Society*, edited by M. Niepokuj, M. van Clay, V. Nikiforidou & D. Feder, Berkeley: University of California, Berkeley, pp. 471–490.
- Prince, A.S. (1983) Relating to the Grid. *Linguistic Inquiry* 14: 19–100.
- Prince, A.S. and P. Smolensky (2004 [1993]) Optimality theory: constraint interaction in generative grammar. Malden, MA: Blackwell. Originally distributed as Rutgers University Center for Cognitive Science Technical report #2.
- Protopapas, A. (2006) On the use and usefulness of stress diacritics in reading Greek. *Reading and Writing* 19: 171–198.

- Protopapas, A., S. Gerakaki and S. Alexandri (2006) Lexical and default stress assignment in reading Greek. *Journal of Research in Reading* 29 (4): 418–432.
- Pruitt, J.S. (1995) The Perception of Hindi Dental and Retroflex Stop Consonants by Native Speakers of Japanese and American English. Ph.D. dissertation, University of South Florida, Tampa.
- Pulgram, E. (1970) *Syllable, Word, Nexus, Cursus*. The Hague: Mouton de Gruyter.
- Pulleyblank, D. (1988) Underspecification, the feature hierarchy and Tiv vowels. *Phonology* 5: 299–326.
- Pullum, G. (1976) The Duke of York Gambit. *Journal of Linguistics* 12: 83–102.
- Pullum, G. and A.M. Zwicky (1988) The syntax-phonology interface, in F.J. Newmeyer (ed.) *Linguistics: The Cambridge Survey (Language: Psychological and Biological Aspects (Volume 3))*. Cambridge: Cambridge University Press, pp. 255–280.
- Pye, S. (1986) Word-final devoicing of obstruents in Russian. *Cambridge Papers in Phonetics and Experimental Linguistics* 5: 1–10.
- Raimy, E. and C. Cairns (ed.) (forthcoming) *Architecture and Representations in Phonological Theory*. Cambridge, MA: MIT Press.
- Ralli, A. (1988) Elements de morphologie du Grec Moderne. Ph.D. dissertation, Université de Montréal.
- Ralli, A. and L. Touratzidis (1992) A computational treatment of stress in Greek inflected forms. *Language and Speech* 35 (4): 435–453.
- Recasens, D. (1987) An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences, *Journal of Phonetics* 15: 299–312.
- Recasens, D., M.D. Pallarès and J. Fontdevila (1998) An electropalatographic and acoustic study of temporal coarticulation for Catalan dark /l/ and German clear /l/. *Phonetica* 55: 53–79.
- Rehg, K.L. (2001) Pohnpeian possessive paradigms: The smart solution, the dumb solution and the Pohnpeian solution, in J. Bradshaw and K.L. Regh (eds) *Issues in Austronesian Morphology. A focusschrift for Byron W. Bender [=Pacific Linguistics, 519]*. Canberra: Research School of Pacific and Asian Studies, The Australian National University, pp. 217–233.
- Reiss, C. (2008) Constraining the learning path without constraints, or the OCP and NoBanana, in B. Vaux and A. Nevins (eds) *Rules, Constraints, and Phonological Phenomena*. Oxford: Oxford University Press, pp. 252–301.
- Rennison, J. (1986) On tridirectional feature systems for vowels, in J. Durand (ed.) *Dependency and Non-Linear Phonology*. London: Croom Helm, pp. 281–303.
- Revithiadou, A. (1999) Headmost Accent Wins. Ph.D. dissertation, University of Leiden. LOT Dissertation Series No. 15. Utrecht: LOT.
- Revithiadou, A. (2006) Prosodic filters on syntax: An interface account of second position clitics. *Lingua* 116: 79–111.
- Revithiadou, A. (2008) A cross-dialectal study of cliticization in Greek, in C. Ewen, H. van der Hulst and N.C. Kula (eds) *Lingua* 118: 1393–1415.
- Revithiadou, A. (to appear) The phonological word, in C. Ewen, B. Hume, K.D. Rice and M. van Oostendorp (eds) *Companion to Phonology*. Oxford: Wiley-Blackwell.
- Revithiadou, A. and V. Spyropoulos (2003) Trapped within a phrase: Effects of syntactic derivation on p-phrasing, in A. Mettouchi and G. Ferré (eds) *Proceedings of IP2003 Prosodic Interfaces*. University of Nantes, pp. 167–172.
- Revithiadou, A. and V. Spyropoulos (2005) The Multiple Spell-Out Hypothesis and the phonological component: Evidence from Greek, in L. Bateman and C. Ussery (eds) *Proceedings of NELS* 35: 523–537.
- Revithiadou, A. and V. Spyropoulos (2006) A typology of Greek clitics with special reference to their diachronic development. Ms., University of the Aegean. [lingBuzz/000496].

- Revithiadou, A. and V. Spyropoulos (2008) Greek object clitic pronouns: A typological survey of their grammatical properties. *Universals and Language Typology. Greek in Typological Perspective* 61: 39–53.
- Revithiadou, A. and V. Spyropoulos (2009) A dynamic approach to the syntax-phonology interface: A case study of Greek, in K. Grohmann (ed.) *InterPhases: Phase-Theoretic Investigations of Linguistic interfaces*. Oxford: Oxford University Press, pp. 202–233.
- Rice, C. and P. Svenonius (1998) Prosodic V2 in Northern Norwegian. Ms., University of Tromsø. [www.hum.uit.no/a/rice/v2/writing/ProsodicV2RiceSvenonius.pdf].
- Rice, K. (1993) A re-examination of the feature [sonorant]: The status of ‘sonorant obstruents’. *Language* 69: 308–344.
- Rice, K. (2009) Whither privative features? An examination of assumptions. Paper presented at the Sixth Old World Conference in Phonology (OCP-6), Edinburgh.
- Rice, K. and P. Avery (1989) On the interaction between sonorancy and voicing. *Toronto Working Papers in Linguistics* 10: 65–82.
- Rice, K. and P. Avery (1991) On the relationship between laterality and coronality, in C. Paradis and J.-F. Prunet (eds) *Phonetics and Phonology (Vol. 2): The Special Status of Coronals; Internal and External Evidence*. New York: Academic Press, pp. 101–124.
- Richards, N. (2006) Beyond Strength and Weakness. Ms., Massachusetts Institute of Technology. [lingBuzz/000325].
- Riehl, A. and A. Cohn (to appear) Partially-nasal segments, in M. van Oostendorp, C.J. Ewen, E. Hume and K. Rice (eds) *The Blackwell Companion to Phonology*. Cambridge, MA: Wiley.
- Riemsdijk, H. van (1978) *A Case Study in Syntactic Markedness*. Peter de Ridder.
- Riggle, J. (2004) Generation, Recognition, and Learning in Finite State Optimality Theory. Ph.D. dissertation, University of California, Los Angeles.
- Ritter, N. (1995) The Role of Universal Grammar in Phonology: A Government Phonology Approach to Hungarian. Ph.D. dissertation, New York University.
- Rivera-Gaxiola, M., G. Csibra, M.H. Johnson and A. Karmiloff-Smith (2000) Electrophysiological correlates of cross-linguistic speech perception in adults. *Behavioural Brain Research* 111: 11–23.
- Rivera-Gaxiola, M., J. Silva-Pereyra and P.K. Kuhl (2005) Brain potentials to native- and non-native speech contrasts in seven- and eleven-month-old American infants. *Developmental Science* 8: 162–172.
- Robins, R.H. (1957 [1970]) Aspects of prosodic analysis, in F.R. Palmer (ed.) *Prosodic Analysis*. London: Oxford University Press, pp. 188–200.
- Roca, I. (1999) Stress in the romance languages, in H. van der Hulst (ed.) *Word Prosodic Systems in the Languages of Europe*. Berlin and New York: Mouton de Gruyter, pp. 659–811.
- Roca, I. and W. Johnson (1999) *A Course in Phonology*. Malden, MA: Blackwell.
- Rochon, E., D. Caplan and G. Waters (1990) Short-term memory processes in patients with apraxia of speech: Implications for the nature and structure of the auditory verbal short-term memory system. *Journal of Neurolinguistics* 5: 237–264.
- Romani, C. and A. Calabrese (1998) Syllabic constraints in the phonological errors of an aphasic patient. *Brain and Language* 64: 83–121.
- Romani, C. and C. Galluzzi (2005) Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cognitive Neuropsychology* 22: 817–850.
- Roosman, L. (2006) Melodic structure in Toba Batak and Betawi Malay word prosody, in V. van Heuven and E. van Zanten (eds) *Prosody in Indonesian Languages*, LOT Occasional Series 9. Utrecht: LOT, pp. 89–115.

- Rose, Y. and K. Demuth (2006) Vowel epenthesis in loanword adaptation: Representational and phonetic considerations. *Lingua* 116: 1112–1139.
- Rose, Y., B. MacWhinney, R. Byrne, G. Hedlund, K. Maddocks, P. O'Brien and T. Wareham (2006) Introducing Phon: A software solution for the study of phonological acquisition, in D. Bamman, T. Magnitskaia and C. Zaller (eds) *Proceedings of the 30th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, pp. 489–500.
- Rosen, S. and P. Iverson (2007) Constructing adequate non-speech analogues: What is special about speech anyway? *Developmental Science* 10: 165–169.
- Ross, J.R. (1967) Constraints on Variables in Syntax. Ph.D. dissertation, Massachusetts Institute of Technology.
- Roussou, A. and I.-M. Tsimpli (2006) On Greek VSO again. *Journal of Linguistics* 42: 317–354.
- Rubach, J. (1986) Abstract vowels in three dimensional phonology: the yers. *The Linguistic Review* 5: 247–280.
- Rubach, J. (1996) Shortening and ambisyllabicity in English. *Phonology* 13: 197–237.
- Rubach, J. (1997) Extrasyllabic consonants in Polish: Derivational Optimality Theory, in I. Roca (ed.) *Derivations and Constraints in Phonology*. Oxford: Clarendon, pp. 551–581.
- Rubach, J. and G. Booij (2003) Lexical phonology – Overview, in W. Frawley (ed.) *International Encyclopedia of Linguistics* (Second edition). Oxford: Oxford University Press, pp. 443–447.
- Rumelhart, D.E. (1989) The architecture of mind: A Connectionist approach, in M. Posner (ed.) *Foundations of Cognitive Science*. Cambridge, MA: MIT Press, pp. 134–159.
- Rumelhart, D.E. and J.L. McClelland (1986) On learning the past tenses of English verbs, in D.D.E. Rumelhart, J. McClelland and the PDP Research Group (eds) *Parallel Distributed Processing* (Volume 2). Cambridge, MA: MIT Press, pp. 216–271.
- Rumelhart, D.E., J.L. McClelland and the PDP Research Group (eds.) (1986) *Parallel Distributed Processing: Exploration in the Micro-Structure of Cognition*. 2 Volumes. Cambridge, MA: MIT Press.
- Russell, K. (1999) MOT: sketch of an OT approach to morphology. Ms., University of Manitoba.
- Saffran, J.R., R.N. Aslin and E.L. Newport (1996b) Statistical learning by 8-Month-old infants. *Science* 274: 1926–1928.
- Saffran, J.R., E.L. Newport, R.N. Aslin (1996a) Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35: 606–621.
- Sag, I. (1974) The Grassmann's Law ordering pseudoparadox. *Linguistic Inquiry* 5: 591–607.
- Sagey, E. (1986) The Representation of Features and Relations in Non-Linear Phonology. Ph.D. dissertation, Massachusetts Institute of Technology.
- Saltzman, E. (1995) Intergestural timing in speech production: Data and modeling. *Proceedings of the XIII International Congress of Phonetic Sciences* 2: 84–91.
- Saltzman, E. and K.G. Munhall (1989) A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1: 333–82.
- Saltzman, E., H. Nam, J. Krivokapić and L.M. Goldstein (2008) A task-dynamic toolkit for modeling the effects of prosodic structure on articulation, in P.A. Barbosa, S. Madureira and C. Reis (eds) *Proceedings of the Speech Prosody 2008 Conference*, Brazil: Campinas, 175–184.
- Salverda, A.P., D. Dahan and J. McQueen (2003) The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition* 90: 51–89.
- Samuel, A.G. (1987) Lexical uniqueness effects on phonemic restoration. *Journal of Memory and Language* 26: 36–56.

- Sandalo, F. and H. Truckenbrodt (2001) Some notes on phonological phrasing in Brazilian Portuguese. *MIT Working Papers in Linguistics* 43: 81–105.
- Sanders, N. (2003) Opacity and Sound Change in the Polish Lexicon. Ph.D. dissertation, University of California, Santa Cruz.
- Sandler, W. (1989) *Phonological Representation of the Sign: Linearity and Non-Linearity in American sign Language*. Dordrecht: Foris Publications.
- Sandler, W. (1993) A sonority cycle in American Sign Language. *Phonology* 10: 243–279.
- Sandler, W. (2008) The syllable in sign language: Considering the other natural language modality, in B. Davis and K. Zajdo (eds) *Ontogeny and Phylogeny of Syllable Organization, Festschrift in Honor of Peter MacNeilage*. New York: Taylor & Francis, pp. 379–408.
- Sandler, W. and M. Aronoff (2007) Is phonology necessary for language?. Paper presented at the Emergence of Language Structures workshop, University of California and San Diego.
- Sandler, W. and D. Lillo-Martin (2006) *Sign Language and Linguistic Universals*. Cambridge, Cambridge University Press.
- Sandler, W., I. Meir, C. Padden and M. Aronoff (2005) The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences* 102: 2661–2665.
- Saussure, F. de (1916 [1959]) *Course in General Linguistics*. New York: Philosophical Society [English edition of *Cours de Linguistique Générale*. Paris: Payot].
- Saussure, F. de (1995 [1916]) *Cours de linguistique générale*. Paris: Payot. Compiled by Charles Bally and Albert S'échehaye. Critical edition by Tullio de Mauro.
- Schadeberg, T. (1987) Silbenanlautgesetze im Bantu. *Afrika und Übersee* 70: 1–17.
- Schane, S.A. (1968) *French Phonology and Morphology*. Cambridge, MA: The MIT Press.
- Schane, S.A. (1974) How abstract is abstract?, in A. Bruck, R.A. Fox, and M.W. LaGaly (eds) *Papers from the Parasession on Natural Phonology*. Chicago, IL: Chicago Linguistic Society, pp. 297–317.
- Schane, S. (1984) The fundamentals of particle phonology. *Phonology Yearbook* 1: 129–155.
- Schane, S. (1995) Diphthongization in particle phonology, in J.A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Oxford: Blackwell, pp. 586–608.
- Scheer, T. (1996) *Une théorie de l'interaction directe entre consonnes. Contribution au modèle syllabique CVCV, alternances e– dans les pré xes chèques, structure interne des consonnes et la théories X-barre en phonologie*. Phd, Université Paris 7.
- Scheer, T. (2004) *A Lateral Theory of Phonology. Vol.1: What is CVCV, and Why Should It Be?* Berlin and New York: Mouton de Gruyter.
- Scheer, T. (2006) How non-phonological information is processes in phonology. Handout of EGG-06 course, Olomouc, July 27 - August 10, 2006.
- Scheer, T. (2008) Why the prosodic hierarchy is a diacritic and why the interface must be direct, in Jutta Hartmann, Veronika Hegedüs and Henk van Riemsdijk (ed.) *Sounds of Silence: Empty Elements in Syntax and Phonology*. Amsterdam: Elsevier, pp. 145–192. [www.unice.fr/dsl/tobias.htm].
- Scheer, T. (in press a) *A Lateral Theory of Phonology. On Locality, Morphology and Phonology in Phonology* (Volume 2). Berlin: Mouton de Gruyter.
- Scheer, T. (in press b) What OT is, and what it is not. Review of *The Cambridge Handbook of Phonology*. P. de Lacy (ed.) *Journal of Linguistics* [appears in 2010]. [www.unice.fr/dsl/tobias.htm].
- Scheer, T. (forthcoming a) How morpho-syntax talks to phonology. A survey of extra-phonological information in phonology since Trubetzkoy's Grenzsignale. Berlin: Mouton de Gruyter.
- Scheer, T. (forthcoming b) Crazy rules, regularity and naturalness, in J. Salmons and P. Honeybone (eds) *The Handbook of Historical Phonology*. Oxford: Oxford University Press.

- Scheer, T. (2010a) Intermodular argumentation: Morpheme-specific phonologies are out of business in a phase-based architecture, in N. Shir and L. Rochman (eds) *The Sound Pattern of Syntax*. Oxford: Oxford University Press, 333–351.
- Scheer, T. (2010b) Review of Gussmann (2007) *The Phonology of Polish*. *Studies in Polish Linguistics* 5: 109–158.
- Scheer, T. and P. Szigetvári (2005) Unified representations for stress and the syllable. *Phonology* 22: 37–75.
- Schiller, N.O., M. Bles and B.M. Jansma (2003) Tracking the time course of phonological encoding in speech production: An event-related brain potential study. *Brain Research. Cognitive Brain Research* 17: 819–831.
- Schuchardt, H. (1885 [1972]) On sound laws: Against the Neogrammarians, in T. Vennemann and T.H. Wilbur (eds) *Schuchardt, The Neogrammarians and the Transformational Theory of Phonological Change*. Frankfurt: Athenäum, 39–72.
- Schwartz, B. and R. Sprouse (1996) L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12: 40–72.
- Scobbie, J.M. (1991a) Attribute Value Phonology. Ph.D. dissertation, University of Edinburgh.
- Scobbie, J.M. (1991b) Towards Declarative Phonology. *Edinburgh Working Papers in Cognitive Science* 7: 1–26.
- Scobbie, J.M. (1993) Constraint Violation and Conflict from the Perspective of Declarative Phonology. *Canadian Journal of Linguistics* 38: 155–168.
- Scobbie, J.M. (1997) *Autosegmental Representation in a Declarative Constraint-Based Framework*. New York: Garland Press.
- Scobbie, J.M., J. Coleman and S. Bird (1996) Key aspects of declarative phonology, in J. Durand and B. Laks (eds) *Current Trends in Phonology: Models and Methods* (Volume 2). Salford, Manchester: ESRI, pp. 685–709.
- Scobbie, J.M., F. Gibbon, W.F. Hardcastle and P. Fletcher (2000) Covert contrast as a stage in the acquisition of phonetics and phonology, in M. Broe and J. Pierrehumbert (eds) *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon*. Cambridge: Cambridge University Press, pp. 194–207.
- Ségéral, P. and T. Scheer (2001) La Coda-Miroir. *Bulletin de la Société de Linguistique de Paris* 96: 107–152.
- Ségéral, P. and T. Scheer (2008) The Coda Mirror, stress and positional parameters, in Carvalho et al. (eds) *Lenition and Fortition*. Berlin and New York: Mouton de Gruyter pp. 483–518.
- Selinker, L. (1972) Interlanguage. *International Review of Applied Linguistics* 10: 209–231.
- Selkirk, E.O. (1972) The Phrase Phonology of English and French. Ph.D. dissertation, Massachusetts Institute of Technology.
- Selkirk, E.O. (1974) French liaison and the X-bar notation. *Linguistic Inquiry* 5: 573–590.
- Selkirk, E.O. (1980a) [1978]. On prosodic structure and its relation to syntactic structure. Bloomington, IN: Indiana University Linguistics Club.
- Selkirk, E.O. (1980b) Prosodic domains in phonology: Sanskrit revisited, in M. Aronoff and M-L. Kean (eds) *Juncture*. Saratoga, CA: Anma Libri, pp. 107–129.
- Selkirk, E.O. (1981a) On prosodic structure and its relation to syntactic structure, in T. Fretheim (ed.) *Nordic Prosody II*. Trondheim: TAPIR, pp. 111–140.
- Selkirk, E.O. (1981b) On the nature of phonological representation, in T. Myers, J. Laver and J. Anderson (eds) *The Cognitive Representation of Speech*. Amsterdam: North Holland, pp. 379–388.
- Selkirk, E.O. (1982) The syllable, in Hulst, H.G. van der and N. Smith (eds) *The Structure of Phonological Representations, Part II*. Dordrecht: Foris Publications, pp. 337–384.

- Selkirk, E.O. (1984a) On the major class features and syllable theory, in M. Aronoff and R.T. Oerhle (eds) *Language and Sound Structure: Studies in Phonology Presented to Morris Halle by His Teacher and Students*. Cambridge, MA: MIT Press, pp. 107–136.
- Selkirk, E.O. (1984b) *Phonology and Syntax: The Relation Between Sound and Structure*. Cambridge, MA: MIT Press.
- Selkirk, E.O. (1986) On derived domains in sentence phonology. *Phonology* 3: 371–405.
- Selkirk, E.O. (1995a) The prosodic structure of function words, in J. Beckman L. Walsh Dickey and S. Urbanczyk (eds) *University of Massachusetts Occasional Papers in Linguistics 18: Papers on Optimality Theory*. University of Massachusetts at Amherst: Graduate Linguistic Student Association, pp. 439–469.
- Selkirk, E.O. (1995b) Sentence prosody: Intonation, stress, and phrasing, in J.A. Goldsmith (ed.) *The Handbook of Phonological Theory*. Cambridge, MA and Oxford: Blackwell, pp. 550–569.
- Selkirk, E.O. (2000) The interaction of constraints on prosodic phrasing, in G. Bruce and M. Horne (eds) *Prosody: Theory and Experiment. Studies Presented to Gösta Bruce*. Dordrecht: Kluwer Academic Publishers, pp. 231–261.
- Selkirk, E.O. (2002) The syntax-phonology interface, in N.J. Smelser and P.B. Baltes (eds) *International Encyclopedia of the Social and Behavioral Sciences*. Oxford: Elsevier Science, pp. 15407–15412.
- Selkirk, E.O. and T. Shen (1990) Prosodic domains in Shanghai Chinese, in S. Inkelas and D. Zec (eds) *The Phonology-Syntax Connection*. Chicago, IL and London: The University of Chicago Press, pp. 313–338.
- Selkirk, E.O. and K. Tateishi (1988) Minor phrase formation in Japanese, in L. Macleod, G. Larson and D. Brentari (eds) *Papers from the Annual Regional Meeting of the Chicago Linguistic Society* 24. Part I. Chicago, IL: CLS, pp. 316–336.
- Selkirk, E.O. and K. Tateishi (1991) Syntax and downstep in Japanese, in C. Georgopoulos and R. Ishihara (eds) *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda*. Dordrecht: Kluwer Academic Publishers, pp. 519–544.
- Semenza, C., G.M.L. Bencini, L. Bertella, I. Mori, R. Pignatti and F. Ceriani (2007) A dedicated neural mechanism for vowel selection: A case of relative vowel deficit sparing the number lexicon. *Neuropsychologia* 45: 425–430.
- Senghas, A., S. Kita and A. Özyürek (2004) Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* 305: 1779–1782.
- Sezer, E. (1983) On non-final stress in Turkish. *Journal of Turkish Studies* 5: 61–69.
- Shallice, T. and E. Warrington (1977) Auditory-verbal short-term memory and conduction aphasia. *Brain and Language* 4: 479–491.
- Shattuck-Hufnagel, S. and A.E. Turk (1996) A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25: 193–247.
- Shepard, R.N., C.I. Hovland and H.N. Jenkins (1961) Learning and memorization of classifications. *Psychological Monographs: General and Applied* 75.13, number 517.
- Shibatani, M. (1973) The role of surface phonetic constraints in generative phonology. *Language* 49 (1): 87–106.
- Shinohara, S. (2004) Emergence of universal grammar in foreign word adaptations, in R. Kager, J. Pater and W. Zonneveld (eds) *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, pp. 292–320.
- Shukla, M., M. Nespore and J. Mehler (2007) An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology* 54: 1–32.
- Shvachkin, N.K. (1973) The development of phonemic speech perception in early childhood, in C. Ferguson and D. Slobin (eds) *Studies of child language development*. New York: Holt, Rinehart and Winston, pp. 91–127, (original work published 1948).

Bibliography

- Siegel, Dorothy (1974) Topics in English Morphology. Ph.D. dissertation, Massachusetts Institute of Technology.
- Sievers, E. (1881) *Grundzüge der Phonetik*. Leipzig: Breitkopf and Hartel.
- Silverman, D. (1992) Multiple scansion in loanword phonology: Evidence from Cantonese. *Phonology* 9: 289–328.
- Silverman, D. (2000) Hypotheses concerning the phonetic and functional origins of tone displacement in Zulu. *Studies in African Linguistics* 29 (2): 3–32.
- Silverman, D. (2006a) *A Critical Introduction to Phonology: Of Sound, Mind, and Body*. London and New York: Continuum.
- Silverman, D. (2006b) The diachrony of labiality in Trique, and the functional relevance of gradience and variation, in L.M. Goldstein, D.H. Whalen and C.T. Best (eds) *Papers in Laboratory Phonology VIII*. Berlin: Mouton de Gruyter, pp. 133–154.
- Silverman, D. (2010) Neutralization and anti-homophony in Korean. *Journal of Linguistics* 46 (2): 453–482.
- Silverman, D. (in prep.) *Neutralization: Rhyme and Reason in Phonology*. Cambridge: Cambridge University Press.
- Sims, A. (in press) Avoidance strategies, periphrasis and paradigmatic competition in Modern Greek, in J.P. Blevins and F. Ackerman (eds) *Periphrasis and Paradigms*. Stanford, CA: CSLI.
- Singh, L., J. Morgan and K. White (2004) Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language* 51: 173–189.
- Skousen, R. (1981) Analogical sources of abstractness, in D. Goyvaerts (ed.) *Phonology in the 80s*. Ghent: Story-Scientia, pp. 55–92.
- Slowiaczek, L.M. and D. Dinnsen (1985) On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13: 325–341.
- Smith, N. and D. Apoussidou (in prep.) Recessive accent in Ancient Greek revisited: The final coda effect.
- Smith, N.S.H. (1988) Consonant place features, in H. van der Hulst and N.S.H. Smith (eds) *Features, Segmental Structure and Harmony Processes (Part I)*. Dordrecht: Foris Publications, pp. 209–235.
- Smith, N.V. (1973) *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press.
- Smolensky, P. (1986) Information processing in dynamical systems: Foundations of Harmony Theory, in D.E. Rumelhart, J.L. McClelland and the PDP Research Group (eds) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Volume 1: Foundations. Cambridge, MA: MIT Press, Bradford Books, pp. 194–281.
- Smolensky, P. (1987) Connectionist AI, symbolic AI, and the brain. *Artificial Intelligence Review* 1: 95–109.
- Smolensky, P. (1988) On the proper treatment of connectionism. *Brain and Behavioural Sciences* 11: 1–74.
- Smolensky, P. (1991) Connectionism, constituency and the language of thought, in B. Loewer and G. Rey (eds) *Meaning in Mind: Fodor and his Critics*. Oxford: Blackwell, pp. 201–227.
- Smolensky, P. (1996) On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27: 720–731.
- Smolensky, P. and G. Legendre (2006) *The Harmonic Mind. From Neural Computation to Optimality-Theoretic Grammar*, 2 Volumes. Cambridge, MA: MIT Press.
- Sohn, H.-S. (1987) Underspecification in Korean Phonology. Ph.D. dissertation, University of Illinois at Urbana-Champaign.
- Sommerstein, A. (1974) On phonetically motivated rules. *Journal of Linguistics* 10: 71–94.

- Sommerstein, A. (1977) *Modern Phonology*, London: Edward Arnold.
- Song, J.Y., M. Sundara and K. Demuth, K. (2009) Effects of phonology on children's production of English 3rd person singular –s. *Journal of Speech, Language, and Hearing Research* 52: 623–642.
- Soto, S., N. Sebastián-Gallés and A. Cutler (2001) Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language* 45: 412–432.
- Southwood, M.H. and A. Chatterjee (1998) Phonological and articulatory disturbances in a case of primary progressive aphasia. *Aphasiology* 12: 161–177.
- Spyropoulos, V. and I. Philippaki-Warbuton (2001) Subject and EPP in Greek: The discontinuous subject hypothesis. *Journal of Greek Linguistics* 2: 149–186.
- Spyropoulos, V. and A. Revithiadou (2009) Subject chains in Greek and PF processing, in C. Halpert, J. Hartman and D. Hill (eds) *Proceedings of the 2007 Workshop in Greek Syntax and Semantics at MIT*. Cambridge, MA: MWPL, pp. 293–309.
- Stager, C.L. and J.F. Werker (1997) Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature* 388: 381–382.
- Stampe, D. (1972a) A Dissertation on Natural Phonology. Ph.D. dissertation, University of Chicago.
- Stampe, D. (1972b) How I Spent my Summer Vacation. Ph.D. dissertation, University of Chicago. Reproduced by the Indiana Linguistics Club in 1979.
- Stanley, R. (1967) Redundancy rules in phonology. *Language* 43 (2): 393–436.
- Steels, L. (2000) Language as a complex adaptive system, in M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J. Julian Merelo, H.-P. Schwefel (eds) *Lecture Notes on Computer Science. Parallel Problem Solving from Nature (PPSN-VI)*. Berlin: Springer, pp. 17–26.
- Steriade, D. (1982) Greek Prosodies and the Nature of Syllabification. Ph.D. dissertation, Massachusetts Institute of Technology.
- Steriade, D. (1988) Greek accent: A case for preserving structure. *Linguistic Inquiry* 19: 271–314.
- Steriade, D. (1993a) Closure, release, and nasal contours, in M.K. Huffman and R. Krakow (eds) *Nasals, Nasality and the Velum*. San Diego, CA: Academic Press, pp. 401–470.
- Steriade, D. (1993b) Positional neutralization. Paper presented at NELS 23, University of Massachusetts, Amherst.
- Steriade, D. (1995a) Positional neutralization. Ms., University of California.
- Steriade, D. (1995b) Underspecification and markedness, in J. Goldsmith (ed.) *Handbook of Phonological Theory*. Oxford: Basil Blackwell, pp. 114–174.
- Steriade, D. (1997) Phonetics in phonology: The case of laryngeal neutralization. Ms., University of California.
- Steriade, D. (1999a) Alternatives to syllable-based accounts of consonantal phonotactics, in O. Fujimura, B.D. Joseph and B. Palek (eds) *Proceedings of the 1998 Linguistics and Phonetics Conference: Item, Order in Language and Speech*. Prague: Karolinum Press, pp. 205–245.
- Steriade, D. (1999b) Licensing laryngeal features. *UCLA Working Papers in Phonetics* 3: 25–146.
- Steriade, D. (1999c) Phonetics in phonology: The case of laryngeal neutralization, in M. Mordon (ed.) *UCLA Working Papers in Linguistics* 2: 25–146.
- Steriade, D. (2001a) Directional asymmetries in place assimilation: A perceptual account, in E. Hume and K. Johnson (eds) *The Role of Speech Perception in Phonology*. San Diego, CA: Academic Press, pp. 219–250.
- Steriade, D. (2001b) The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Unpublished manuscript. University of California, Los Angeles.

- Steriade, D. (2007) Contrast, in P. de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, pp. 139–157.
- Stevens, K.M. (1998) *Acoustic Phonetics*. Cambridge, MA: MIT Press.
- Stevens, K.N. and D.H. Klatt (1974) Role of formant transitions in the voiced-voiceless distinction for stops. *Journal of Acoustical Society of America* 55: 653–659.
- Stillings, N., S. Weisler, C. Chase, M. Feinstein, J. Garfield and E. Rissland (1995) *Cognitive Science. An Introduction*. Cambridge, MA: MIT Press.
- Stockwell, R. and J. Bowen (1965) *The Sound Patterns of English and Spanish*. Chicago, IL: University of Chicago Press.
- Stoel-Gammon, C. (1985) Phonetic inventories, 15–24 months: A longitudinal study. *Journal of Speech and Hearing Research* 28: 505–512.
- Stoel-Gammon, C. and A.V. Sosa (2006) Phonological development, in E. Hoff and M. Schatz (eds) *Blackwell Handbook of Language Development*. Hoboken, NJ: John Wiley and Sons, pp. 238–256.
- Stokoe, W. (1960) Sign language structure: An outline of the visual communication systems of the American deaf. *Studies in Linguistics: Occasional Papers*. Buffalo, NY: University of Buffalo.
- Storkel, H.L. (2002) Restructuring similarity neighborhoods in the developing mental lexicon. *Journal of Child Language* 29: 251–274.
- Storkel, H.L. (2004) Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research* 47: 1454–1468.
- Straka, G. (1963) La division de sons du langage en voyelles et consonnes peut-elle être justifiée? *Travaux de Ling. et de Littérature* (Université de Strasbourg) 1: 17–99.
- Strange, W. (ed.) (1995a) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore, MD: York Press.
- Strange, W. (1995b) Cross-language studies of speech perception: A historical review, in W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium, MD: York Press, pp. 3–45.
- Strange, W. and S. Dittmann (1984) Effects of discrimination training on the perception of /r-/l/ by Japanese adults learning English. *Perception and Psychophysics* 36: 131–145.
- Streeter, L.A. (1976) Language perception of two-month-old infants shows effects of both innate mechanisms and experience. *Nature* 259: 39–41.
- Studdert-Kennedy, M. and L. Goldstein (2003) Launching language: The gestural origin of discrete infinity, in M.H. Christiansen and S. Kirby (eds) *Language Evolution: The States of the Art*. Oxford: Oxford University Press, 235–254.
- Sundara, M., K. Demuth and P.K. Kuhl (2008, Nov.) Positional effects on 2-year-olds' comprehension and production of 3rd person singular –s. Poster presented at the 33rd Boston University Conference on Language Development, Boston, Massachusetts.
- Sutton-Spence, R. and B. Woll (1999) *The Linguistics of British Sign Language: An Introduction*. Cambridge: Cambridge University Press.
- Swadesh, M., and C.F. Voegelin (1939) A problem in phonological alternation. *Language* 15: 1–10.
- Swingle, D. and R.N. Aslin (2000) Spoken word recognition and lexical representation in very young children. *Cognition* 76: 147–166.
- Swingle, D. and R.N. Aslin (2002) Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science* 13: 480–484.
- Szigetvári, P. (2002) Dismantling syllable structure. *Acta Linguistica Hungarica* 48: 155–181.
- Szigetvári, P. (2006) The markedness of the unmarked. *Acta Linguistica Hungarica* 53: 433–447.

- Szigetvári, P. (2007) Branching onsets and syncope in English. *Language Sciences* 29: 408–425.
- Szigetvári, P. (2008) Two directions for lenition, in Carvalho et al. (eds) *Lenition and Fortition*, Berlin and New York: Mouton de Gruyter pp. 561–592.
- Szigetvári, P. and T. Scheer (2005) Unified representations for the syllable and stress. *Phonology* 22: 37–75.
- Tabain, M. (2003) Effects of prosodic boundary on /aC/ sequences: Articulatory results. *Journal of the Acoustical Society of America* 113: 2834–2849.
- Tabain, M. and P. Perrier (2005) Articulation and acoustics of /i/ in preboundary position in French. *Journal of Phonetics* 33: 77–100.
- Takahashi, T. (1993) A farewell to constituency. *UCL Working Papers in Linguistics* 5: 375–410.
- Takahashi, T. (1999) Constraint interaction in Aranda stress, in S.J. Hannahs and M. Davenport (eds) *Issues in Phonological Structure: Papers from an International Workshop*. Amsterdam and Philadelphia, PA: John Benjamins, Volume 196 of *Amsterdam Studies in the Theory and History of Linguistic Science*, pp. 151–181.
- Taylor, J. (2002) *Cognitive Grammar*. Oxford: Oxford University Press.
- Tesar, B. (1998) An iterative strategy for language learning. *Lingua* 104: 131–145.
- Tesar, B. (2004) Contrast analysis in phonological learning. [ROA#695].
- Tesar, B. (2009) Learning phonological grammars for output-driven maps. To appear in *The Proceedings of NELS* 39. [ROA#1013].
- Tesar, B. and P. Smolensky (2000) *Learnability in Optimality Theory*. Cambridge: MIT Press.
- Thevenin, D.M., R.E. Eilers, D.K. Oller and L. Lavoie (1985) Where's the drift in babbling drift? A cross-linguistic study. *Applied Psycholinguistics* 6: 3–15.
- Thompson, R., D. Vinson and G. Vigliocco (2009) The link between form and meaning in American Sign Language: Lexical processing effects. *Journal of Experimental Psychology: Learning, Memory and Cognition* 35: 550–557.
- Tincoff, R. and P.W. Jusczyk (1999) Some beginnings of word comprehension in 6-month-olds. *Psychological Science* 10: 172–175.
- Törkenczy, M. (1994) Coronal underspecification and the *CCiVCi constraint. *The Even Yearbook* 1: 225–232.
- Tourville, J. (1991) Licensing and the Representation of Floating Nasals. Ph.D. dissertation, McGill University.
- Tranel, B. (1981) *Concreteness in Generative Phonology. Evidence from French*. Berkeley, CA: University of California Press.
- Treiman, R. and C. Danis (1988) Syllabification of intervocalic consonants. *Journal of Memory and Language* 27: 87–104.
- Trigo, L. (1993) The inherent structure of nasal segments, in M.K. Huffman and R. Krakow (eds) *Nasals, Nasality and the Velum*. San Diego, CA: Academic Press, pp. 369–400.
- Trinh, T. (2009) A constraint on copy deletion. *Theoretical Linguistics* 35: 183–227.
- Trofimovich, P. and W. Baker (2006) Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition* 28: 1–30.
- Trost, J.E. and G.J. Canter (1974) Apraxia of speech in patients with Broca's aphasia: A study of phoneme production accuracy and error patterns. *Brain and Language* 1: 63–79.
- Trouvain, J. and U. Gut (eds) (2007) *Non-Native Prosody: Phonetic Description and Teaching Practice*. Berlin: Mouton de Gruyter.
- Trubetzkoy, N.S. (1939) *Grundzüge der Phonologie*. Göttingen: Vandenhoeck and Ruprecht [Translated in 1969 by C. Baltaxe as *Principles of Phonology*. Berkeley/Los Angeles, CA: University of California Press].

- Truckenbrodt, H. (1995) Phonological Phrases: Their Relation to Syntax, Focus, and Prominence. Ph.D. dissertation, Massachusetts Institute of Technology.
- Truckenbrodt, H. (1999) On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30: 219–255.
- Turing, A. (1936–1937) On computable numbers, with an application to the Entscheidungs problem. *Proceedings of the London Mathematical Society*, series 2, 42: 230–265.
- Turk, A.E. and S. Shattuck-Hufnagel (2000) Word-boundary-related durational patterns in English. *Journal of Phonetics* 28: 397–440.
- Tversky, A. (1977) Features of similarity. *Psychological Review* 84: 327–352.
- Tversky, A. and I. Gati (1978) Studies of similarity, in E. Rosch and B.B. Lloyd (eds) *Cognition and Categorization*. New York: John Wiley and Sons, pp. 79–98.
- Tversky, A. and I. Gati (1982) Similarity, separability, and the triangle inequality. *Psychological Review* 89: 123–154.
- Twaddell, W.F. (1935) *On Defining the Phoneme*, [=Language Monograph, 16]. Baltimore, MD: Linguistic Society of America, University of Pennsylvania. Reprinted in Joos (1957), 55–80.
- Tyler, M.D. and A. Cutler (2009) Cross-language differences in cue use for segmentation. *Journal of the Acoustical Society of America* 126: 367–376.
- Tyler, M.D., L. Tyler and D.K. Burnham (2005) The delayed trigger voice key: An improved analogue voice key for psycholinguistic research. *Behavior Research Methods* 37: 139–147.
- Uffmann, C. (2007), Intrusive [r] and optimal epenthetic consonants. *Language Sciences* 29: 451–476.
- Uriagereka J. (1999) Multiple spell-out, in S.D. Epstein and N. Hornstein (eds) *Working Minimalism*. Cambridge, MA: MIT Press, pp. 251–282.
- Uyechi, L. (1996) *The Geometry of Visual Phonology*. Stanford, CA: CSLI.
- Valdois, S. (1990) Internal structure of two consonant clusters, in J.-L. Nespoulous and P. Villiard (eds) *Morphology, Phonology and Aphasia*. New York: Springer, pp. 253–269.
- Varley, R. and S.P. Whiteside (2001) What is the underlying impairment in acquired apraxia of speech?. *Aphasiology* 15: 39–49.
- Vaux, B. (1998) *The Phonology of Armenian*. Oxford: Clarendon Press.
- Vaux, B. (2003) Syllabification in armenian, universal grammar, and the lexicon. *Linguistic Inquiry* 34: 91–125.
- Vaux, B. (2008) Why the phonological component must be serial and rule-based, in B. Vaux and A. Nevins (eds) *Rules, Constraints, and Phonological Phenomena*. Oxford: Oxford University Press, pp. 20–60.
- Vennemann, T. (1971) Natural generative phonology. Paper presented at the Annual Meeting of the Linguistic Society of America, St. Louis.
- Vennemann, T. (1972a) On the theory of syllabic phonology. *Linguistische Berichte* 18: 1–18.
- Vennemann, T. (1972b) Phonological uniqueness in natural generative grammar. *Glossa* 6:105–116.
- Vennemann, T. (1972c) Sound change and markedness theory: On the history of the German consonant system, in R.P. Stockwell and R.K.S. Macaulay (eds) *Linguistic Change and Generative Theory. Essays from the UCLA Conference on Historical Linguistics in the Perspective of Transformational Theory (1969)*. Bloomington, IN: Indiana University Press, pp. 230–274.
- Vennemann, T. (1974a) Phonological concreteness in natural generative grammar, in R. Shuy and C. Bailey (eds) *Toward Tomorrow's Linguistics*. Washington, DC: Georgetown University Press, pp. 202–219.

- Vennemann, T. (1974b) Words and syllables in natural generative grammar, in A. Bruck, R. Fox and M. La Galy (eds) *Papers from the Parasession on Natural Phonology*. Chicago, IL: Chicago Linguistic Society, pp. 346–374.
- Vennemann, T. (1976[1971]) Vowel alternations in English, German and Gothic: Remarks on realism in phonology, in M.A. Jazayery, E.C. Polomé and W. Winter (eds) *Linguistic and Literary Studies in Honor of Archibald A. Hill, vol. I: General and Theoretical Linguistics*. Lisse: de Ridder, pp. 337–359.
- Ventureyra, V., C. Pallier and H-Y. Yoo (2004) The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics* 17: 79–91.
- Vihman, M.M., R.A. DePaolis and T. Keren-Portnoy (2009) Babbling and words: A dynamic systems perspective on phonological development, in E.L. Bavin (ed.) *The Cambridge Handbook of Child Language*. Cambridge: Cambridge University Press, pp. 163–182.
- Visch-Brink, E.G. (1999) Words in Action: Retrieval Errors in Aphasia, a Topic for Therapy. Ph.D. dissertation, Erasmus University Rotterdam.
- Vitevitch, M.S. and P.A. Luce (1998) When words compete: Levels of processing in perception of spoken words. *Psychological Science* 9: 325–329.
- Vogel, I. (1985) On constraining prosodic rules, in H. van der Hulst and N. Smith (eds) *Advances in Nonlinear Phonology*. Dordrecht: Foris Publications, pp. 217–233.
- Vogel, I. (1986) External sandhi rules operating between sentences, in H. Andersen (ed.) *Sandhi Phenomena in the Languages of Europe*. Berlin: Mouton de Gruyter, pp. 55–64.
- Vogel, I. (1991) Level ordering in Italian Lexical Phonology? in P.M. Bertinetto, M. Kenstowicz and M. Loporcaro (eds) *Certamen Phonologicum II*. Torino: Sellier and Rosenberg, pp. 81–101.
- Vogel, I. (1997) Prosodic phonology, in M. Maiden and M. Parry (eds) *The Dialects of Italy*. London and New York: Routledge, pp. 58–67.
- Vogel, I. (1999) Subminimal constituents in prosodic phonology, in S.J. Hannahs and M. Davenport (eds) *Issues in Phonological Structure. Papers from an International Workshop*. Amsterdam/Philadelphia, PA: John Benjamins, pp. 249–267.
- Vogel, I. (2009) The status of the clitic group, in J. Grijzenhout and B. Kabak (eds) *Phonological Domains: Universals and Deviations*. Berlin: Mouton de Gruyter, pp. 15–46.
- Vouloumanos, A. and J.F. Werker (2004) Tuned to the signal: The privileged status of speech for young infants. *Developmental Science* 7: 270–276.
- Vouloumanos, A. and J.F. Werker (2007) Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science* 10: 159–164.
- Walker, R. and G.K. Pullum (1999) Impossible segments. *Language* 75: 764–780.
- Walley, A.C. (2005) Speech perception in childhood, in D.B. Pisoni and R.E. Remez (eds) *Handbook of Speech Perception*. Oxford: Blackwell, pp. 449–468.
- Wang, C. (1995) The Acquisition of English Word-Final Obstruents by Chinese Speakers. Ph.D. dissertation, State University of New York, Stony Brook.
- Wang, Q. (2008) Perception of English Stress by Mandarin Chinese Learners of English: An Acoustic Study. Ph.D. dissertation, University of Victoria.
- Wang, W. (1968) Vowel features, paired variables and the English vowel shift. *Language* 44: 694–708.
- Wang, Y., A. Jongman and J. Sereno (2006) Second language acquisition and processing of Mandarin tone, in P. Li, L.H. Tan, E. Bates and O.J.L. Tzeng (eds) *Handbook of East Asian psycholinguistics* (Volume 1: Chinese). Cambridge: Cambridge University Press, pp. 250–257.
- Wängler, H.H. (1958) *Atlas Deutscher Sprachlaute*. Berlin: Akademie-Verlag.
- Warner, N., A. Jongman, J. Sereno and R. Kems (2004) Incomplete neutralization and other sub-phonemic differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32: 251–276.

- Warner, N., T. Otake and A. Takayuki (to appear) Intonational structure as a word boundary cue in Japanese. *Language and Speech*.
- Warrington, E.K. and T. Shallice (1969) The selective impairment of auditory verbal short-term memory. *Brain* 92: 885–896.
- Watkins, K.E., A.P. Strafella and T. Paus (2003) Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41: 989–994.
- Wayland, R.P., S.G. Guion, D. Landfair and B. Li (2006) Native Thai speakers' acquisition of English word stress patterns. *Journal of Psycholinguistic Research* 35: 285–304.
- Wedel, A.B. (2004) Self-Organization and Categorical Behavior in Phonology. Ph.D. dissertation, University of California, Santa Cruz.
- Wedel, A.B. (2006) Exemplar models, evolution and language change. *The Linguistic Review* 23: 247–274.
- Weijer, J.M. van de (1996) *Segmental structure and complex segments*. Tübingen: Niemeyer.
- Welby, P. (2007) The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication* 49: 28–48.
- Wells, J.C. (1982) *Accents of English*. Cambridge: Cambridge University Press.
- Wells, J.C. (1990) *Longman Pronunciation Dictionary*. Harlow: Longman.
- Wells, R.S. (1949) Automatic alternations. *Language* 25: 99–116.
- Werker, J. and L. Polka (1993) Developmental change in speech perception: New challenges and new directions. *Journal of Phonetics* 21: 83–101.
- Werker, J.F. and J.S. Logan (1985) Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37: 35–44.
- Werker, J.F. and R.C. Tees (1983) Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology* 37: 278–286.
- Werker, J.F. and R.C. Tees (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7: 49–63.
- Werker, J.F. and S. Curtin (2005) PRIMIR: A developmental framework of infant speech Processing. *Language Learning and Development* 1: 197–234.
- Werker, J.F., J.H.V. Gilbert, K. Humphrey and R.C. Tees (1981) Developmental aspects of cross-language speech perception. *Child Development* 52: 349–355.
- Wernicke, C. (1874 [1977]) *Der aphasische symptomenscomplex: eine psychologische studie auf anatomischer basis* [Translated 1977 by G.H. Eggert as 'The aphasia symptom complex. A psychological study on an anatomic basis', in *Wernicke's Works on Aphasia: A Sourcebook and Review*. New York: Mouton, pp. 91–145.
- Wheeldon, L.R. and W.J.M. Levelt (1995) Monitoring the time course of phonological encoding. *Journal of Memory and Language* 34: 311–334.
- White, K.S. and J. Morgan (2008) Sub-segmental detail in early lexical representations. *Journal of Memory and Language* 59: 114–132.
- White, L. (1989) *Universal Grammar and Second Language Acquisition*. Amsterdam/Philadelphia, PA: John Benjamins.
- Wiese, R. (1996) *Phonology of German*. Oxford: Oxford University Press.
- Wightman, C.W., S. Shattuck-Hufnagel, M. Ostendorf and P.J. Price (1992) Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91: 1707–1717.
- Wilbur, R. (1973) The Phonology of Reduplication. Ph.D. dissertation, University of Illinois at Urbana-Champaign.
- Wilbur, R. (1987) *American Sign Language: Linguistic and Applied Dimensions*. Boston, MA: College-Hill Press.
- Williams, E. (1976) Underlying tone in Margi and Igbo. *Linguistic Inquiry* 7: 463–484.
- Williams, E. (2007) Dumping Lexicalism, in G. Ramchand and C. Reiss (eds) *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press, pp. 353–381.

- Wilshire, C.E. and J.-L. Nespoulous (1997) The role of syllabic frequency in aphasic phonological errors. *Brain and Language* 60: 147–150.
- Wilshire, C.E. and J.-L. Nespoulous (2003) Syllables as units in speech production: Data from aphasia. *Brain and Language* 84: 424–447.
- Wilson, C. (2001) Consonant cluster neutralisation and targeted constraints. *Phonology* 18: 147–197.
- Woll, B. (2001) The sign that dares to speak its name: Echo phonology in British sign language, in P. Boyes-Braem and R. Sutton-Spence (eds) *The Hands are the Head of the Mouth*. Hamburg: Signum, pp. 87–98.
- Wonderly, W.W. (1951a) Zoque I: Introduction and bibliography. *International Journal of American Linguistics* 17: 1–9.
- Wonderly, W.W. (1951b) Zoque II: Phonemes and morphophonemes. *International Journal of American Linguistics* 17: 105–123.
- Wonderly, W.W. (1951c) Zoque III: Morphological classes, affix list, and verbs. *International Journal of American Linguistics* 17: 137–162.
- Wonderly, W.W. (1951d) Zoque IV: Auxiliaries and nouns. *International Journal of American Linguistics* 17: 235–251.
- Wood, S. and W.J. Hardcastle (2000) Instrumentation in the assessment and therapy of motor speech disorders: A summary of techniques and case studies with EPG, in I. Papathanasiou (ed.) *Acquired Neurogenic Communication Disorders*. London: Whurr, pp. 203–248.
- Wright, H.H. and R.J. Shisler (2005) Working memory in aphasia: Theory, measures, and clinical implications. *American Journal of Speech-Language Pathology* 14: 107–118.
- Wright, R. (2001) Perceptual cues in contrast maintenance, in E. Hume and K. Johnson (eds) *The Role of Speech Perception in Phonology*. San Diego, CA: Academic Press, pp. 251–277.
- Yip, M. (1988) The OCP and phonological rules: A loss of identity. *Linguistic Inquiry* 19: 65–100.
- Yip, M. (1998) Identity avoidance in phonology and morphology, in S. Lapointe, D. Brentari and P. Farrell (eds) *Morphology and its Relation to Phonology and Syntax*. Stanford, CA: CSLI Publications, pp. 216–246.
- Yip, M. (2003) Casting doubt on the Onset-Rime distinction. *Lingua* 113: 779–816.
- Yip, M. (2006) The symbiosis between perception and grammar. *Lingua* 116 (7): 950–975.
- Yoshida, S. (1993) Licensing of empty nuclei: The case of Palestinian vowel harmony. *The Linguistic Review* 10: 127–159.
- Younger, B. and L.B. Cohen (1986) Developmental change in infants' perception of correlations among attributes. *Child Development* 57: 803–815.
- Youssef, A. and I. Mazurkewich (1998) The acquisition of English metrical parameters and syllable structure by adult native speakers of Egyptian Arabic (Cairene dialect), in S. Flynn, G. Martohardjono and W. O'Neil (eds) *The Generative Study of Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 303–332.
- Yu, A.C.L. (2004) Explaining final obstruent voicing in Lezgian: Phonetics and history. *Language* 80: 73–97.
- Yu, A.C.L. (2007) Understanding near mergers: The case of morphological tone in Cantonese. *Phonology* 24: 187–214.
- Zamuner, T.S. (2003) *Input-Based Phonological Acquisition*. New York: Routledge.
- Zamuner, T.S. (2009) Phonotactic probabilities at the onset of language development: Speech production and word position. *Journal of Speech, Language, and Hearing Research* 52: 49–60.
- Zamuner, T.S. and D. Ohala (1999) Preliterate children's syllabification of intervocalic consonants, in A. Greenhill, H. Litterfield and C. Tano (eds) *Proceedings of the 23rd*

Bibliography

- Boston University Conference on Language Development* (Volume 2). Somerville, MA: Cascadilla Press, pp. 753–763.
- Zamuner, T.S., L.A. Gerken and M. Hammond, M. (2004) Phonotactic probabilities in young children's speech production. *Journal of Child Language* 31: 515–536.
- Zec, D. (1988) Sonority Constraints on Prosodic Structure. Ph.D. dissertation, Stanford University.
- Zec, D. (1993) Rule domains and phonological change, in S. Hargus and E.M. Kaisse (eds) *Studies in Lexical Phonology*. San Diego, CA: Academic Press, pp. 365–405.
- Zec, D. (1995) Sonority constraints on syllable structure. *Phonology* 12: 85–129.
- Zec, D. and S. Inkelas (1990) Prosodically constrained syntax, in S. Inkelas and D. Zec (eds) *The Phonology-Syntax Connection*. Chicago, IL and London: University of Chicago Press, pp. 365–378.
- Zec, D. and S. Inkelas (1991) The place of clitics in the Prosodic Hierarchy, in D. Bates (ed.) *Proceedings of the Tenth West Coast Conference on Formal Linguistics (WCCFL)*. Stanford, CA: Stanford Linguistics Association, pp. 505–519.
- Zsiga, E.C. (1995) An acoustic and electropalatographic study of lexical and postlexical palatalization in American English, in B. Connell and A. Arvaniti (eds) *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence*. Cambridge: Cambridge University Press pp. 282–302.
- Zwicky, A. (1974) Homing in: On arguing for remote representations. *Journal of Linguistics* 10: 55–69.

Language Index

- Adamorobe Sign Language 275
Al-Sayyid Bedouin Sign Language 255, 275
American Sign Language (ASL) 255, 257–8, 268–9
Arabic 314
 Cairene Arabic 75, 93
 Classical Arabic 152
 Egyptian Arabic 311
 Jordanian Arabic 311
 Moroccan Arabic 83–4
 Palestinian Arabic 94
Armenian, New Julfa 150
Barra Gaelic 378
Berber (Tashlhiyt) 197
British Sign Language (BSL) 255, 260–4, 266–7, 273
Bulgarian 97
Cama (Ebrié) 57
Catalan 305, 388, 390
Central Sierra Miwok 98
Chimwiini 237
Chinese 304, 312, 314–15
 Cantonese Chinese 300, 307
 Mandarin Chinese 307
 Shanghai Chinese 238
 Xiamen Chinese 237
Czech 72, 312, 315
Dakota 97
Dutch 23, 89, 93, 98, 218–21, 279, 333–4, 388
 Sign Language 255, 272
English 33, 72–3, 83, 96, 98, 147, 153, 176, 178–9, 191, 193, 198, 225, 232, 279, 281, 304, 315–16, 321, 327, 344, 349, 354–5, 386, 390
 American English 77, 344, 389
 Canadian English 138, 144, 169, 424
 Middle English 45, 389
 Modern English 76
 Old English 406
 Scottish Standard English 77
 Singapore English 151
 Standard British English 77
 Standard Southern British English 66, 77
 Tyneside English 390
Filipino 282
Finnish 97, 99, 161, 313
French 64, 77–8, 83, 98, 184, 227–8, 284, 301, 312, 314–15, 349–50, 363, 417, 421
 Canadian French 146, 167, 304, 311
 Sign Language 274, 276
 Standard French 149
German 83, 96, 98, 100, 136, 174–5, 179, 215, 218–21, 312, 315, 370, 372–3
Greek 100–1, 233, 245, 249
 Athenian Greek 230
 Ancient Greek 69, 101, 105, 405
 Modern Greek 95, 100–1, 106
Hausa 240–1, 251
Hawaiian 188
Hungarian 75, 83, 311, 313–14

Language Index

- Italian 78, 233, 236, 239, 312, 315
 Sicilian Italian 229
 Standard Italian 90
 Tuscan Italian 90, 229
- Japanese 47, 98, 71, 279, 304–9, 350
 Yamato Japanese 47
- Javanese Indonesian 98
- Kera 9, 11–12
- Kikuyu 281
 Kinyambo 98
- Korean 188, 284, 308–10, 316, 350,
 354, 364
- Kwakwala 75
- Lardil (Pama-Nyungan) 75, 151–2
- Latin 83, 100, 385
- Lithuanian 75, 142, 153, 157
- Luganda 48
- Macedonian 97
- Malay
 Betawi Malay 98
 Sri Lankan Malay 98
- Manam 71
- Mohawk (Northern Iroquoian) 145,
 155, 168
- Mongolian, Khalka 75
- Navajo 49
- Norwegian 141, 156
- Nupe (Kwa; Central Nigeria) 148
- Nzebi (Bantu; Gabon) 143
- Pacific Yupik 100
- Palauan (Malayo-Polynesian) 147
- Pintupi 99
- Polish 83, 97, 311–13, 315, 388, 421
- Proto-Indo-European 43
- Puerto Rican Spanish (Boston) 383
- Russian 46, 97, 184, 312, 315, 371–3,
 386, 388
- Shona 387
- Sign Language of the Netherlands 255,
 272
- Somali 83
- Southern Barasano 57, 58
- Spanish 51, 53, 100, 284, 307, 314–16,
 380, 383
 Andalusian Spanish 388
 Caribbean Spanish 390
- Swahili 387
- Swedish 285, 386
- Taiwanese 350, 354
- Tangale (Chadic; Nigeria) 143, 157,
 162, 167
- Thai 304, 316
- Turkish 97, 152, 306, 314
- Tuyuca 58
- Ukrainian 373
- Xhosa 329
- Yawelmani 144, 155, 179
 Yokuts Yawelmani 144, 155
- Yiddish 70, 80
- Yucatec Mayan Sign Language 275
- Yupik Eskimo 98
- Zoque 50, 55
- Zulu 302

Subject Index

- abstractness 140, 147, 398, 408, 410, 413, 418
- accent 96, 105–6, 351, 353, 355, 357
- acoustics 43, 44, 203–5, 211
 - acoustic analyses 4, 26
 - acoustic data 7
 - acoustic information 3
 - acoustic measurements 12
 - acoustic phonetics 6
 - acoustic salience 282, 285, 291
- algorithm 235
 - arboreal mapping algorithm 241
 - end-based algorithm 243, 248
 - relation-based algorithm 235
- alignment 238–9
 - alignment constraints 197
- allomorphy 103–4, 147, 166, 406–7, 442
 - allophone 136, 155, 170
 - allophonic level 138
- alternation 152, 166, 370, 372–3, 407–8, 411
 - Alternation Condition 148, 180, 409–10
- ambisyllabicity 89
- analogy 166, 170, 412, 442
- analytic morphology 68
 - non-analytic morphology 68
- Anticipatory Eye Movement Paradigm 18, 20–2
- anti-cyclicity 436
- anti-derivationalism 422, 425, 427, 436–7
- aphasia 321–2, 324, 326, 329
 - Broca's aphasia 323, 326, 328
 - Conduction aphasia 323
 - phonemic jargon aphasia 334
 - Wernicke's aphasia 323, 326, 335
- apraxia of speech (AOS) 322–4, 340
- arboreal mapping 241
- archiphoneme 136
- areal features 14
- articulation 204, 217
 - articulator features 304
 - articulatory gesture 36, 168
 - articulatory planning 321, 323, 329, 333
- Articulatory Organ Hypothesis 286, 295
- Articulatory Phonology 356–7, aspiration 218–21
- Autosegmental phonology 141, 149, 182–3, 29
 - autosegmental licensing 87
 - autosegmental representations 412–15, 418
 - autosegmental tier 270, 416
 - autosegments 415, 417
- babbling 17
 - babbling drift 280
 - canonical babbling 279
- behaviourism 422, 432–3
 - behavioural methodologies 18, 21
- Bidirectional Phonetics and Phonology 212, 216, 221
- bilinguals 21
- binarity 240
 - binary branching 79, 88
 - binary features 34, 402, 418
 - bivalent features 34

- bleeding 150, 153–4
 - absolute bleeding 149
 - counter-bleeding 150, 153–4, 164
 - mutual bleeding 152
- bonding strength 360, 367
- boundary 22–8, 235, 240, 415, 420
 - boundary marking 348–9, 351–2
 - boundary signals 377–9, 391–2
- brain 321, 325–6
 - brain lesions 321 324
 - brain-style computation 424, 429
- branching 232, 236, 239–40
- broad transcription 26
- c-command 230, 233
- C-to-C government 84–5, 87–8
- candidate chain 155, 159
- categorical perception 258–9, 280–1, 283, 434
- categorization 20
- chain shift 143
- CHILDES database 24
- clicks 302
- closed syllable shortening 78, 79
- cluster reduction 26, 328, 333
- coarticulation 206, 386–7
- coda capture 88, 89
- coda constraints 89
- cognitive development 286–7, 296
- Cognitive Phonology 140
- Cognitive Science 398, 402, 422, 424, 429, 433, 443
- Comparative Markedness 439
- competence 429–30
- complexity 44, 46
 - complex onset 306–7
 - compound expression 44
- computation 433–4, 438, 442
 - computational economy 430
 - computational trope (OT) 435
- conduite d'approche 322, 327–8, 335
- connectionism 422, 424, 429, 432
- consonant clusters 66, 69, 82–5, 306–7, 309, 310
 - strong cluster 66
 - weak cluster 66
- consonantal contact 309,
- conspiracy 150, 420, 427
- constituent 233–4, 237
 - prosodic constituent 231, 233, 235, 239, 243
 - syntactic constituent 226, 233–5, 229
- constraints 108–9, 116, 157, 159,
 - chapter 7, 217, 219, 427–8, 430–1, 437
 - constraint-and-repair approach 413
 - cue constraints 217, 219, 221
 - derivational constraint 150
 - weighted constraints 124, 130,
- containment, Principle of 139
- context 155–6
 - context-free rule 141, 155
- contrast 5, 8, 12, 25, 300–2, 304–5
 - covert contrasts 26
 - liquid contrasts 304
- convergence 155, 158, 160
- co-occurrence restriction 47, 53, 60
- co-phonologies 399, 436
- Copy Theory of Movement 244–8
- coronal 304
- corpus data 18, 23–4, 28
- correspondence 165, 195–6, 198
 - Correspondence Theory 139, 165, 195
 - diagonal correspondence 151
 - transderivational
 - correspondence 165
- counterbleeding 141, 145–6, 152
 - non-opaque counterbleeding 152
 - overapplication 152, 165
- counterfeeding 141, 143, 151
 - surface-true counterfeeding 151
 - underapplication 150–1
- cross sectional designs 16
- cue 217
 - perceptual cue 211, 213
- culminativity 98
- CVCV theory 441 *see also* Strict CV Theory
- Cyclic Phonology 164
 - cyclic rules 163
 - cyclic derivation 401, 442–3
 - cyclicity 162, 401, 436
- data collection 3
- Declarative Phonology 139, 184, 186, 200, 420, 425–6

- deletion 332
- Dependency Phonology 34, 35, 419, 441
 dependency 41, 44, 45, 46
 dependent 45, 46, 48, 56
 head-dependency relation 45, 48, 55
- dependent measure 22
- Derivational Optimality Theory (DOT) 437, 439
- derivationalism 422, 424–5, 436–7
 derivations 138ff.
- derived environment 163–4
 derived environment effects 408–11
- descriptive approach 3, 6, 14
- destructive alteration 138
- determinant 154
- devoicing 300
- diachronics 406, 408
- diphthongs 71, 86–7
- direct access hypothesis 227, 230
- Direct Realism 216
- discrimination paradigms 21
- disjunctive ordering 161
- distinctive features 36–7, 39, 136, 149
- distractors 20
- distributed computation 16
- documentation 4
- domain c-command 230, 242
 prolific domains 243
 prosodic domains 236–7
 spell-out domain 248
- domain-general 286–7, 296
- domain-initial strengthening 350,
 353–4, 362, 364–5
- dominance constraint 263–4
- downstep 10, 11
- duality of patterning 255, 275
- Duke-of-York derivation 145–6, 156,
 167
- duplication problem 182, 198
- duration 12, 96
- dysarthria 324
- echo phonology 262
 echoism 28
 echolalia 28
- effort minimization 210
- Electromyogram (EMG) 351
- Electropalatography (EPG) 344, 349–50
- Element Theory 34, 42
 elements 34, 42, 204
- elicitation tasks 23, 28
- elision 65–6
- Elsewhere Condition 140, 161, 182, 198–9
- Emergence of the Unmarked, The (TETU) 192
- empiricism 422, 432–4, 443
- Empty Category Principle (ECP) 422
- empty data cells 26
- encapsulation 399
- endangered languages 4, 8
- enriched inputs 439
- environment 141, 150
- epenthetic vowels 12
- ethical permission 5
- EVAL 155, 158–9, 192–5, 428, 439
- evaluation measure 411
 evaluation metric 410
- excrescence 168
- Exemplar-based Model 361, 367
 exemplar modeling 380–3
 Exemplar Theory 208–9
- experiments (perception,
 psycholinguistic) 13, 15
- explicit judgments 19
- extrametricality 117, 119–20
- extrinsic rule ordering 140, 161–2, 166,
 408, 422–3, 425–6, 440
- factorial typology 190, 207, 429, 438
- faithfulness 112, 116
 faithfulness constraints 112, 116, 188,
 191, 196
- familiarity preference 18
- featurally underspecified lexicon 291, 293
- feature 36–8, 204–5, 220–1, 262, 269–72,
 283, 291–3
- feature-changing 138–9
- Feature Geometry 38, 141, 149, 182,
 270, 319, 418–19
 class features 40
 feature matrix 402
 Feature Theory 36
 feature tree 3
 major class features 39, 40
 Sound Pattern of English (SPE)
 features 39, 181, 269

- feeding 150–1, 153
 - cross-derivational feeding 152
 - non-gratuitous feeding 152
 - self-destructive feeding 152
- feet 99, 116, 420
 - bounded feet 99
 - foot structure 11, 12, 102, 118,
 - Iambic feet 100
 - Iambic-Trochaic Law 100
 - unbounded feet 99
- fieldwork 3, 15
- filler items 2
- filter 6
- final devoicing 175, 186, 196,
204–5, 221
- fingerspelling 266, 268
- first words 2
- Firthian Prosodic Analysis 139
 - Firthian school 415
- floating segment 416, 421
- focus 141, 143
- free generation-cum-filters 431
- Freedom of Analysis 194
- frequency 375–6, 380, 392
- functional approaches 209–10, 212,
214, 217
 - functional load 376, 379, 390
- fundamental frequency 96

- Garde's Principle 388
- geminate 89
 - geminate integrity 89
 - geminate-singleton contrasts 305
 - virtual geminates 89
- GEN 155, 158, 194–5, 428, 436
- gesture 46, 211, 220
- glides 72–3, 207
- global conditions 427, 430
 - global rules 150
- Government and Binding 241
- Government Phonology 34–5, 167,
153–4, 184–5, 187, 200, 419–20, 425,
439, 441
 - C-to-C government 84–5, 87–8
 - government 84–5
 - Proper-government 83
 - V-to-V government 84, 87–8

- gradience 323, 327
 - gradient representations 206
- gradualness 155
- grid theory 99
 - grid marks 99, 417

- handshape 258–9, 263–4, 268
 - hand-internal movement 259, 271
 - hand orientation 262, 270–1
 - Hand Tier model 270–1
- Harmonic Grammar 107–8, 112, 427
 - harmonic bounding 154, 159,
190, 194
 - harmonic improvement 155, 157–9
 - Harmonic Phonology 140, 427
 - harmonic score 109
 - Harmonic Serialism 155, 158–9,
426, 439
- Head 48, 58
- Head-Driven Phrase Structure
Grammars (HPSC) 426
- Headturn Preference Procedure
(HPP) 19
- heavy 100, 105
 - weight 124
- hiatus 86–7, 198
 - hiatus resolution 199
- history of generative phonology 397,
403, 418, 404
- homorganic consonant 33, 33
 - hormoganicity 36, 41
 - homorganic nasal 33, 36

- iconicity 256–7, 259, 271, 275
- imitation 10
 - imitation task 10
- indexed constraints 399, 436
- input-orientation 154
- instrumental analysis 6
- intensity 96
- interactionism 442–3
- interface 202–4, 242, 244, 248
 - interface constraints 436
 - interface requirements 430
 - translation between modules 398,
400
- interlanguage 300, 307, 317, 319

- intermediate representation 141, 147, 156
- Intermodal Preferential Looking Procedure (IPLP) 19
- intonation 346–7, 364
- invariance 202–3
- inverted T model 399–400, 433
- juncture phonemes 400
- Kazan school 370, 391
- Kinematics 357–9
- Laboratory Phonology 203, 208, 342, 345–6
- language acquisition 43, 166, 278, 280, 284
- language contact 265
- language exposure 281, 283–4
- language games 6, 9
- lateralization 309
- learnability 95, 101, 111–12, 410
- length 96, 100, 106
- levels of representation 135
- lexical 339
 - anti-lexicalism 412, 444
 - Lexical Category Condition 238
 - lexical contrast 35, 46
 - lexical development 18
 - lexical diffusion 18, 27
 - lexical encoding 303, 305
 - lexical frequency 203, 336
 - Lexical Phonology 164, 399–400, 410, 412, 414, 418, 439, 443
 - lexical representation 20, 288–9, 296, 331
 - lexical retrieval 257
 - Lexical Segmentation 364–6
 - lexical stress 97, 117
- lexicon 291, 293–6, 410, 421, 438
- liaison 65–6, 421
- licensing 86–7, 90
 - Licensing Constraints 425, 438
 - V-to-V licensing 87,
- linearization 227, 242–3, 245–6
 - linearization of chains 244, 246
- literal paraphasia 321–2, 335, 339
- local movement 259
- location 259, 261–3, 270–2
- logorrhea 322
- longitudinal studies 16, 22
- loudness 98
- Lyman's Law 47
- mapping 211, 214–15, 222, 233, 235, 237, 242
 - arboreal mapping 241
 - end-based mapping 235, 237–8
 - relation-based mapping 235
- markedness 189, 196–8, 207, 264–5, 270, 299–300, 307, 317, 404
 - markedness constraints 197, 217–18
 - Markedness Differential Hypothesis 307
 - markedness hierarchies 197
- marking conventions 181
- Maximal Application Hypothesis 149
- Maximum Onset Principle (MOP) 183, 191
- Meinhof's Law 48
- melodic primes 419
- merge 247
- metathesis 142–3, 157
- metrical grid 417, 420
- Metrical Phonology 141, 149
- minimal pairs 137–8, 169, 259, 262
- Minimal Sonority Distance 306
- minimalism 400–1, 430
 - Minimalist Program 241–2
- mirror neurons 216
- misapplication 150
- mispronunciation 2
- modality 255, 257, 259
 - modality-independence 400
- modular 323, 329
 - modularity 216, 222, 327, 398, 436, 443
- monovalent 34
 - single-valued 34
- mora 204
- moraic theory 416
- morpheme 147
 - Morpheme Structure Condition (MSC) 141, 170, 177–8

Subject Index

- morpho-phonology 405, 407
 morphophonemic level 136–8
morpho-syntax 399–400, 436, 405, 420
Motor Theory of Speech
 Perception 215–16
Move-alpha 422, 427, 431
Move-Hold model 268–70
 movement 258–9, 261–2, 268,
 270–3, 275
MP-rules 407
multi-linear representation 413, 415
multiple opacity 155
- nasal place assimilation 41
nasality 34, 47, 50
 nasalization 50, 54, 55, 58, 149, 309
native-language bias 26
Native Language Magnet Model 303
native speaker intuitions 67
natural class 33
Natural Generative Phonology 153,
 181, 406, 412, 418
Natural Phonology 181, 183, 200, 206,
 398, 406–7, 414
 naturalness 149, 181, 186, 205
naturalistic data 23
near-infrared spectroscopy (NIRS) 19
near-merger 388–9
neighbourhood density 28, 295–6
neologism 329, 335
neural commitment 285, 295
neutralization 36, 52, 136, 148, 206
 absolute neutralization 148–9, 409–10
 contextual neutralization 148
 near-neutralization 388, 390
NoBanana 428
nominalism 136
non-linear representations 413, 415–16
non-manual features 256, 262
non-word repetition tasks 27
 non-word stimuli 28
 nonce words 9, 11
novelty preference 2
nucleus-coda interaction 4
- Obligatory Contour Principle
 (OCP) 183, 417, 420, 422, 443
- on-line measure 22
onset maximization 69, 89
OO (output-output) faithfulness 436
opacity 150, 155–6, 200, 439–40
Optimality Theory 98, 107, 139, 154,
 188, 199–200, 207, 300, 307, 425
 with Candidate Chains
 (OT-CC) 439–40
 Stratal Optimality Theory 164–5,
 437, 439
- ordered rules 422, 425, 436, 440
ordering paradoxes 430
orientation 262, 270–1
output-orientated 140, 154, 427–8
overapplication 150–2, 165, 170
overgeneralization errors 25
overgeneration 404, 407–8, 413–14, 429
- P-rules 407–8
Paradigmatic Contrast
 Enhancement 352–4
parallel computation 427, 437, 439, 443
 Parallel Distributed Processing
 (PDP) 422, 434
Particle Phonology 419
path movement 259, 271
pause-insertion task 23
perception 13, 14, 16, 18, 19, 43,
 213–15, 218–19, 308, 313–14,
 316–17, 348, 361–2
 Perceptual Assimilation Model 285,
 295, 302
 perceptual epenthesis 309–10
 perceptual errors 18
 perceptual magnet effect 303
 perceptual magnets 285
 perceptual similarity 302
 production-perception
 asymmetry 305
performance 424, 429
phases 242–3, 248–50, 430
 Phase Impenetrability Condition
 (PIC) 411
phi-function 426
phoneme 136–7, 259, 261, 268–9
 phonemic level 137
 autonomous phonemic level 137

- phonetics 343, 345, 367
 intrinsic phonetic content 137
 phonetic detail 17, 18, 20
 phonetic exposure 302
 phonetic grounding 203–4, 206–8,
 210
 phonetic implementation 202,
 214–15, 220
 phonetic level 135–6
 phonetic similarity 302
 phonetic transcription 5, 26
 Phonetics-Phonology Interface 343,
 345, *Chapter 8*
 Phonetics-Prosody Interface 346–8,
 355
 phonological class 205
 phonological development 280,
 286–7, 292
 phonological parsers 310
 phonological primes 34, 35, 42
 phonological specificity 290, 295
 phonological word 231–2
 phonotactics 306, 308–9
 phonotactic constraints 66, 76, 79
 phonotactic probabilities 28
 phrase 227, 231–2, 239
 intonational phrase 243
 major phrase 238
 phonological phrase 231
 phonological phrasing 231
 phrase-final lengthening 364
 physical realism 136
 Pi-gesture 358–60
 picture naming task 19, 24, 25
 piloting 25
 pitch 96, 105
 place constraint 263–4
 positional faithfulness 196
 Possible-Word Constraint 272–4
 post-lexical 334, 339
 post-lexical phonology 230, 233
 post-nasal voicing 47, 51, 61
 precedence 415–16
 prenasalization 58, 59
 prenasalised stops 52, 58, 59, 60
 Principles and Parameters
 Theory 242
 principum divisionis 136
 privative 34, 35
 privativity 35
 probability matching 383, 386–90
 production 18, 29
 production errors 18, 28
 prominence 95, 347–8, 351–2, 354,
 355–6, 361
 prominence marking 347–8,
 351–3, 355
 proper bracketing 89
 proper government 83, 167
 Proper Inclusion Precedence 161
 prosody 95–6, 280, 290, 362–3,
 366–8, 378–9
 prosodic boundary 352,
 358, 360
 prosodic cues 364
 prosodic domain 361
 Prosodic Hierarchy 222, 231, 347,
 350, 417, 420
 prosodic licensing 86–7
 Prosodic Model 271
 Prosodic Morphology 423
 Prosodic Phonology 231, 234,
 399–401, 420, 436–7
 prosodic strengthening 222, 349,
 351–2, 355–6, 362
 prosodic structure 347–8, 352,
 362–3, 366–7
 Prosodic Structure Hypothesis 227,
 233
 prosodic word 231–2, 237
 prosodically-driven phonetic
 detail 361, 367
 Prosody Analyser 363, 366–7
 Prosody Generator 361–2,
 366–7
 protophones 279, 297
 psychological realism 136

 Radical CV Phonology 34
 ranking paradox 154–5
 rationalism 422, 434
 reaction time measures 22, 26
 real-word stimuli 20, 27, 28
 recency effects 375

Subject Index

- recording (data) 7, 13
 - audio and video 24
 - recording equipment 3, 7
 - MP3 recordings 7
 - WAV files 7
- recoverability 156
- recursion 441
- reductionism 432
- redundancy rules 177–8
- reinforcer 21
- relexicalization 169–70
- repair strategy 308, 310
- representation 139, 168, 434ff.
 - intermediate representation 141, 147
 - non-linear representation 141, 148
 - surface representation 138, 140, 142
 - underlying representation 138, 142, 147, 156, 170
- resyllabification 89
- retroflex 210, 217
- Revised Alternation Condition 409–10
- rhotic 207
- rhyme 75–6, 79
- rhythm 232, 414, 417
 - rhythm rule 232
- Richness of the Base 193–4
- rule-based phonology 139–41, 144, 156
 - rule/list fallacy 433, 438
 - rules-cum-constraints 431
- rules 138, 141, 161, 177–81, 209, 227–8, 230, 233, 407, 422, 428,
 - cyclic rules 162–3
 - global rules 150
 - lexical rules 164, 169
 - opaque rules 150, 169–70
 - post-lexical rules 164
 - rewrite rules 140
 - rule conspiracies 176
 - rule interaction 141, 144–5, 150–1, 154
 - rule sandwiching 145–6, 155–6, 168
 - transient rules 140
 - transparent rules 166–7
- rule ordering 140
 - extrinsic rule ordering 140, 161
 - intrinsic rule ordering 161
 - natural rule ordering 149–50
 - random sequential rule
 - ordering 152–3
 - simultaneous rule application 152–3, 154
 - unnatural rule ordering 149
- sandhi 230
 - sandhi rule 227–8
- scrambling trope (OT) 434, 436–8
- segment 9, 12, 204, 206, 211, 269
 - segmentation 272, 274
- selected fingers constraint 264, 266, 270
- Selection Models 281
- Sequential Markedness Principle 307
- serialism 423–5, 431, 439,
- short term memory 324, 333
- Sign Language 400
- sign-spotting task 10
- skeleton 416, 420–1
 - skeletal slots 416
- slip of the tongue errors 29, 261
 - slip of the ear 29
 - slip of the hand errors 261, 271
- sonority 64, 70, 306–7
 - sonority hierarchies 70, 73
 - sonority sequencing principle 71
- sound change 206
 - diachronic linguistics 207
- spectral structure 96
- speech bias 279
- speech comprehension *see* perception
- Speech Learning Model 300, 303
- speech perception *see* perception
- speech plan 322
- speech sounds 33, 34, 38
- Spell-Out 242, 249–50
 - Multiple Spell-Out 243
 - Multiple Spell-Out Hypothesis 242
- statistical analysis 8
 - statistical significance 13
- Stratal OT 437, 439
 - stratum 164
- stray erasure 68
- stress 311–15, 317, 346–7, 353–4, 363, 414, 417, 420
 - analogy 316
 - deafness 311, 313–14, 317

- quantity sensitive/insensitive
 stress 314, 319
 stress-bearing element 67
 stress shift 121–2, 362
 stress-timed language 354
 Stress Typology Model 314–15
 Strict CV theory 185, 16
 Strict Cycle Condition 182, 410–11, 422
 strict domination 194
 Strict Layer Hypothesis (SLH) 417, 420
 structural change 141, 421, 428
 structural description 141, 153–4, 410,
 428–9
 structuralism 136–7, 402, 434, 443
 American Structuralism 136
 structure preservation 169, 182, 198
 structure-building 139
 structure-checking 139
 substitution 321, 326, 332
 substitution errors 26, 327, 329, 331
 suppletion 147, 166, 406, 442
 suprasegmental features 347, 363
 surface representation 135, 138, 140
 Surface Structure Constraints (SSCs) 179
 syllable 54, 64ff., 211, 221, 268, 273–4,
 306, 310, 312, 315
 degenerate syllable 80–1
 syllable internal organization 74–5
 syllable weight 75–6
 syllabic consonant 73
 universal syllable 70, 80
 vowelless syllable 79–80
 syllable structure 325, 327–8, 330,
 332–3, 16
 symbolic representations 432
 symmetry constraint 263
 Sympathy Theory 155, 439
 syntagmatic contrast enhancement 352,
 354
 syntax-first model 233
 syntax-phonology interface 9

 target 17, 20
 targeted constraints 439
 task dynamics 356, 358
 terminal node 304
 ternary branching 87

 Theory of Constraints and Repair
 Strategies (TCRS) 183–5, 200
 repair strategies 184, 200
 time course analyses 22
 toddlers 19
 tone 9, 10, 11, 14, 315
 tone frame 1
 too many repairs problem 185
 transcription *see* phonetic transcription
 narrow transcription 26
 transducer 212
 transitional probabilities 379, 391–2
 trisyllabic shortening 409, 411, 424
 laxing 409
 trisyllabic window 98, 105, 113,
 117, 124
 trochees 99, 120
 truncation 24, 27
 turbidity 440
 turing machine 422, 443
 Two-Level Morphology 140

 underapplication 150–1
 underlying form 405, 409, 411
 underlying representation 138, 142,
 147–8, 164
 underspecification 419
 Underspecification Theory 171
 underspecified 20
 Uniformitarianism 389
 Universal Grammar 303, 308,
 318, 423
 usage-based approaches 433

 variation 19, 203
 velar softening 424–5
 violation *see* Optimality Theory
 visual expectation paradigm 2
 voice key 26
 Voice Onset Time (VOT) 21, 29, 280,
 344, 350, 354–5
 voicing 47, 50, 51, 205
 contrast 300, 304
 post-nasal voicing 61
 post-sonorant voicing 143, 167
 pre-sonorant voicing 143, 158
 vowel duration 12

Subject Index

vowel harmony 414
vowel quality 96
vowel-zero alternation 83–4, 421

well-formedness conditions 413–14,
423, 427, 431, 443
Window Model 360–1, 367

word 231–2
word frequency 24, 27, 28
word production 17
word stress 98
word-games 23
word-spotting task 273
Wug task 23