INTRODUCTION TO THE HISTORY OF THE ENGLISH LANGUAGE

Written by László Kristó

TABLE OF CONTENTS

INTRODUCTION

The aim of the present textbook is to introduce BA students of English to the basics of the history of English. It is "unorthodox" in a couple of ways. First, it does not discuss the history of the English language in a chronological order, unlike most other textbooks: instead, it starts with a presentation of Early Modern English, which is followed by a discussion of the most important changes that have taken place since then. The reason for this is that students at a BA level, unless wishing to study the early history of English, will most probably not read Middle English – let alone Old English – texts in the original, but they will come across Early Modern English texts (for example, ones by Shakespeare). Therefore, I find it useful to put more emphasis on this period than the earlier ones.

Second, following from the previous point, this textbook discusses earlier periods in the history of English in a less detailed way than usual. I do not believe that a detailed discussion of West Germanic sound changes or a near-complete overview of Old English grammar (just to give some examples) is actually necessary for the average BA student. Instead, emphasis will be laid on those points – such as ablaut, to give an example – which are of some relevance for the understanding of some peculiarities or regularities of present-day English. This does not mean that I consider early periods unimportant: I merely wish to say that they are much less relevant for the target audience. Anyone who is interested in these periods can take further, specialized, courses, or go on to study them at an MA (or even PhD) level. There is a list of suggested readings at the end of each chapter to help those who are more deeply interested in the given topic.

Third, following from the fact that this book is primarily written for students at Pázmány Péter Catholic University, where the subject is taught together with another, namely, an introduction to English dialectology, I have attempted to connect the contents of this textbook to the dialectology course as much as possible, which can hopefully be seen especially in Chapters 2 and 3, where the developments of the past four centuries are discussed.

Finally, I hope that this textbook will be useful for BA students – not only at Pázmány Péter Catholic University, but anywhere else where professors might choose to use it as a set text.

NOTES ON PHONETIC SYMBOLS USED IN THIS BOOK

Generally, English words are transcribed according to the usual conventions in Hungary, that is, following the most widespread dictionaries (Longman or Oxford, for example) and textbooks (such as Nádasdy (2007) or Balogné and Szentgyörgyi (2007)). Some other symbols will, however, be used as well. These are listed below.

Vowels

o like Hungarian *o* in *rop*
oː like Hungarian *ó* in *tó*
ɛ like Hungarian *e* in *per*
ɛː like Hungarian *e* in *per,* but long
eː like Hungarian *é* in *szép*
a like Hungarian *á* in *ráz,* but short
aː like Hungarian *á* in *ráz*

Consonants

x a voiceless velar fricative, like Hungarian *h* in *doh*, or German *ch* in *Bach*
ɣ the voiced counterpart of **x**, as in Spanish *lago*
b^{h} an aspirated **b**
d^{h} an aspirated **d**
g^{h} an aspirated **g**
k^{w} a **k** pronounced with lip rounding (similar to **kw** in *queen,*but one sound)
g^{w} a **g** pronounced with lip rounding (similar to **gw** in *Gwen,*but one sound)
g^{hw} an aspirated **g^{w}**

1 Language change and historical linguistics

1.1 Language history and its study

As everything else in the world, languages, too, change over time. The most easily noticeable aspect of change is in vocabulary: new words are born almost day by day, or old ones acquire new meanings. Think of words such as *facebook*, for instance, which came into existence a couple of years ago, or *net*, which is an ancient word with a new meaning (= 'internet'). You may also find, though examples are more rare, that words which were used earlier have become old-fashioned or even obsolete, such as the pronoun *thou*, which was used in the sense of 'you', but originally only to address one person (like Hungarian *te*).

However, the pronounciation and the grammar of languages also change, but much more slowly, so such changes are more difficult to spot, at least for the non-linguist. An example is provided in (1) below:

(1) *Everyone has a car of his own.*
Everyone has a car of their own.

The first sentence shows the older use, the masculine pronoun *his* being used to refer back to *everyone*. Nowadays, however, *their* is the usual choice – obviously, because of gender correctness, *his* is avoided, although older speakers might still use *his*.

As for pronunciation, changes are even more difficult to spot, though not impossible. In standard British English pronunciation, called Received Pronunciation (RP), words such as *sore, boar, story* were pronounced with a diphthong **ɔə** in the early 20th century. Nowadays, this sounds old-fashioned, the diphthong having been replaced by a long monophthong, i.e. **ɔː**.

The branch of linguistics which studies language change is called **historical linguistics**. It has basically two aspects. First, it deals with language change in general: how and possibly why languages change. It describes the mechanisms of language change and attempts to discover the common types of change in all human languages. This aspect can be called theoretical. On the other hand, historical linguistics also studies the history of individual languages: this aspect can be called applied. Of course, the two aspects are not independent of each other: you can hardly make generalizations about language change unless you study the history of individual languages; on the other hand, to explain the developments found in a given language, you will need to use theoretical linguistic methods. It is beyond the scope of the present textbook to give you a detailed outline of theoretical historical linguistic issues, although some theoretical notions will have to be used. If you are interested more deeply in the subject, you will find some suggested reading at the end of this chapter.

1.2 Internal and external history

The history of a language can be described from two different points of view: **internal** and **external**. Roughly, internal history is the description of changes in the given language: how the pronunciation, the grammar and the vocabulary have changed over time. External history is concerned with the non-linguistic circumstances under which the language has developed: this includes social, cultural or political events that affected the people who speak the language. The examples given in the previous section illustrate internal history: the loss of a

pronoun, or a change from a diphthong to a monophthong are purely linguistic facts. On the other hand, it is part of the external history of the English language, for instance, that it was carried overseas after the discovery of America. The colonization of North America by England was by no means a linguistic event! This does not mean, of course, that external factors – though not linguistic ones by themselves – have no relevance for the internal history of the language. After colonization had begun, the English language gradually started to develop in differing ways in England and in America. By the end of the 18th century, the differences became significant enough to enable us to talk about "British" as opposed to "American" English. A note of warning is needed here, however. Non-linguist tend to think that British English is the same as "Shakespeare's English", which became "corrupted" in North America. This is far from the truth: language is constantly changing, and Britain is no exception, so present-day British English is as different from early 17th century English as present-day American is. (Indeed, in terms of pronunciation, for example, American English is in most ways closer to Shakespeare's pronunciation!) We will have ample occasion to see the details later on.

1.3 The periodization of the history of languages

It is customary in all historical sciences – social, political, biological, etc. – to divide the history of the studied object into periods. This is called **periodization**. In political history, for example, we talk about Antiquity (Hungarian *ókor*) which is usually taken to end in 476 AD, the time when the West Roman Empire collapsed; this period, then, is followed by the Middle Ages, and so on. Of course, as shown by this particular example, periodization is mostly somewhat arbitrary, and the end of one period (and the beginning of the next one) is usually connected to a symbolic date. Things rarely change overnight (except maybe in the case of revolutions), and the social and cultural differences that distinguish Antiquity from the Middle Ages arose quite slowly, taking centuries. The point is that you have to draw the line somewhere, and in this case, 476 AD seems to be a convenient symbolic point.

Similarly, the history of languages is divided into periods, and the periodization generally follows important external historical dates or events. The history of the English language is generally divided into the following periods:

1. **Pre-Old English**, dating from 449 AD, when, according to tradition, three Germanic tribes, the Angles, the Saxons and the Jutes arrived in Britain from their original homeland in southern Scandinavia. (We will discuss this issue later on in more detail.) However, it is only during the 8th century that our first English texts appear (see below), so from the period between 449[1] and 700 we have no direct (written) evidence for the development of English: this is why we call this period Pre-Old English.

2. **Old English**, dating from around 700, because it is from the 8th century that we possess the first surviving English texts. It lasts till about 1100, but opinions vary: some scholars date the end of the Old English period to 1066, when the Norman Conquest took place, while some others date it to the mid-12th century. For the sake of simplicity, I regard 1100 as the end of the Old English period.

[1] From this point on, I omit the abbreviation *AD* after years, since it is self-evident.

3. **Middle English**, from about 1100 to around 1500. Again, opinions vary as to the end of this period (some give the date 1476, when Caxton set up the first printing press in England). For simplicity's sake, I use 1500.

4. **Early Modern English**, from about 1500 to about 1800. Again, some give the symbolic date of 1776, the year of the American Declaration of Independence, as the end date. I will use 1800.

5. **Later Modern English**, from 1800 up to the present.

Finally, it is usual to refer to current English as **Present-day English**. I will only use this term when I wish to emphasize the difference between Later Modern English (say, of the 19th century) and the current situation.

It is useful to point out that I have often chosen the "round numbers", such as 1100 or 1500, for the sake of simplicity. Round numbers are easier to remember, and also recall that language does not change overnight, as it were. For example, the Norman Conquest of 1066 did have some influence on English, especially on vocabulary, but also on pronunciation; however, this does not mean that speakers of English woke up one day to find out that they spoke another language (i.e. Middle English). Instead, the linguistic effects of the Norman Conquest were gradual, taking place over several generations' time, and that is another reason why 1100 has been chosen here as a convenient dividing line between Old and Middle English.

Another point to note is that the history of English did not, strictly speaking, begin in the year 449. It is merely the case that it was at around that time that English started to develop as a language different from its closest relatives. All natural languages have a long history behind them, and English is no exception: indeed, as we will see, its history can be traced back to at least six thousand years – except that before the middle of the 5th century AD, it does not really make any sense to talk about "English" as an individual language, essentially different from its closest relatives (such as Dutch or German, for example). Such issues will be taken up in detail later on.

1.4 The chief types of linguistic change at various levels

As I mentioned earlier, this book does not aim at being a general introduction to the theory of language change; nonetheless, it is inevitable to introduce you to the most elementary notions. This section, then, is far from being exhaustive, and only issues that will be of some relevance in later chapters will be discussed. First, we discuss **lexical change**, i.e. changes in vocabulary. Second, we look at **semantic change**, that is, changes in the meaning of words or grammatical constructions. Third and fourth, we examine **morphological** and **syntactic change**, respectively; and finally, we discuss **phonological change**, that is, changes in the pronounced form of language.

It must be emphasized that these aspects of change are often interconnected, and, indeed, they are sometimes difficult to be treated separately. A change in one part of language may not only go hand in hand with a change in another part, but it may itself cause another change to take place. I hope the discussion below will make this point of interrelatedness reasonably clear.

1.4.1 Lexical change

As mentioned, this aspect of language change is the most easily detectable one: words come and go – some get first old-fashioned, then obsolete, while new words are born, almost day by day. The number of new words created far outnumbers the number of words that die out (so it is completely unneccessary to worry, as many people do, about the loss of a particular word, regarded by them as a proof that language is becoming poorer). Let us now start with loss, then, followed by gain.

During the course of time, words may be lost from the language. **Lexical loss** is hardly ever sudden, of course: first, a given word is more and more rarely used, so that over a couple of generations' time it becomes first **old-fashioned** (used by older speakers only), then it turns **archaic** (i.e. not used by anyone at all, unless for specific purposes, but still understood at least by educated people), after which it may become completely **obsolete**, i.e. dead: it is neither used nor understood by any speaker, not even an educated one (unless he or she has had special training in the history of the language). Let me illustrate these stages in lexical loss now. I will provide some Hungarian examples, too, for the majority of the readers of this book will probably be native speakers of Hungarian, and they might find some examples from their own language helpful.

1. An example of an old-fashioned word is the Auxiliary *shall* with a future meaning, as in *I shall write to her right away*. Nowadays, in natural speech as well as in informal written language, the Auxiliary *will* would be used instead. Such a use of *shall* then, is either old-fashioned, or quite formal (the two often go hand in hand: formal usage tends to be more conservative). Examples of lexical content words (i.e. not grammatical function words) which are now quite old-fashioned (at least for most speakers of English) include *sneakers* 'sports shoes' or *record player*, for instance. From Hungarian, we might think of words such as *távbeszélő* 'telephone' – which was still used in formal written language when I was a young child, but no longer used (not even in formal styles), at least as much as I can observe.

2. Archaic words include ones which are still understood but no longer used. In British English, for example, the word *wireless* 'radio' could be cited – although it might be argued that this word is simply old-fashioned, rather than archaic. Indeed, it is often difficult to draw a clear boundary: what some speakers might find old-fashioned may be considered definitely archaic by others. The pronoun *thou*, on the other hand, is a clear example: it is understood (as a synonym of *you*) by all educated speakers of English, but no one would ever use it nowadays (except for special, such as poetic, purposes). A Hungarian example might be *kend* 'you', or *vala* 'was' – no longer used in standard Hungarian but understood by most speakers.

3. Obsolete words include ones such as *eek* 'also', *to wend* 'to go, to turn' (its past form *went* is still alive, though, being used as the past tense form of *go*), or *soothe* 'true'. If you ever come across a word such as these, you will probably need an etymological dictionary (a dictionary showing the origin of words) to find out its meaning. A Hungarian example could be *héon* 'only', found in the so-called *Funeral Speech* (*Halotti beszéd*), the first surviving complete text in Hungarian.

There are, however, far more new words added to a language than those being lost. Let us now look at the chief types of lexical enrichment.

1. The most frequent instance of the birth of new words is via **word formation**, either by **compounding** (such as *facebook, cellphone*, etc.) or **affixation** (such as *demote, replay, holiness, priceless*; the first two exemplify prefixation, while the latter two illustrate suffixation). Word formation is quite productive, so much so that it is often used by speakers rather intuitively, even unconsciously. For example, there is a suffix *–like* in English, added to nouns to form adjectives, as exemplified by the word *catlike* – meaning 'resembling a cat'. You may freely add this suffix to virtually any noun to express the same idea. For example, if you wish to say that someone looks like Jabba, the vile, disgusting monster of the Star Wars

films, you can say that the person is question is *Jabba-like*. (You may use it in Hungarian, too, saying that somebody is *Jabbaszerű*.)

2. Another instance of lexical enrichment is when a language **borrows** words from another. Such words are called **loanwords** (or **loans** for short). English, for example, has borrowed quite many words from a variety of languages, especially from French, Latin, Scandinavian (= North Germanic, roughly, Danish and Norwegian) and Greek. Examples include *chair, dance, rule, machine* (from French), *interrogate, separate, quorum* (from Latin), *take, law, skirt* (from Scandinavian), or *geology, atom, astronomy* (from Greek). Other languages have also contributed to the enrichment of the English word stock, though to a lesser extent – examples include Italian (e.g. *cello, pizza*), Spanish (e.g. *macho, tortilla*), Arabic (e.g. *algebra, algorithm*), but even Hungarian (e.g. *hussar*, *goulash*, or *coach* – from Hungarian *kocsi*).

3. Words may also come into existence by **abbreviation** (also called **clipping**), whereby a longish word is "cut short". Examples include *ad* (from *advertisement*), *telly* (from *television*) or *math(s)* (from *mathematics*).

4. **Back formation** is also found. A classical example is the verb *to edit*, being back-formed from *editor*. Speakers of English interpreted the word *editor* as 'someone who edits' – in other words, the suffix *–or* was cut off from the noun, yielding the verb *to edit*. Such back formations are the result of speakers feeling that there is a suffix, where originally, there had been none. In this particular case, the origin of this back formation is due to analogy, based on verb-noun pairs such as *bake – baker* (note that the suffixes spelt *-or* and *-er* are both pronounced the same, i.e. **ər**), where the noun denotes the person performing the act. Another example is the noun *pea*, originally *peas*: here, the *–s* was interpreted as a plural marker (quite logically, since the vegetable in question is usually consumed in great numbers), and the singular form *pea* was created. To take a Hungarian example, the original form of present-day *tulipán* 'tulip' was *tulipánt* (still found in this form in the late 18th century). Yet, speakers felt the *–t* to be an accusative marker, so it was removed – as a result, the original nominative form is now an accusative.

5. **Acronyms** are also sources of new words. They arise when the initial letters of phrases are used. Examples include *EU* (for *European Union*) or *USA* (for *United States of America*). In these examples, the initial letters are pronounced separately, according to their alphabetic value. Often, however, acronyms are read out as complete words, exemplified by items such as *NATO* (for *North Atlantic Treaty Organization*), pronounced **ˈneɪtəʊ**. Sometimes, acronyms become fully independent words, that is, they are no longer recognized as acronyms as far as their origin is concerned. An example is the word *laser*, originally an acronym of *Light Amplification by Stimulated Emission of Radiation* – but no one (unless a specialist) is aware of this fact any longer.

6. **Proper names** may also become sources of new words, i.e. common nouns. This happens especially in the case of brand names. The word *hoover*, for example, was originally a brand name of a make of vacuum cleaners; now, at least in British English, it means 'vacuum cleaner' of any make in general. In Hungarian, the word *mirelit*, now meaning 'frozen food' in general, was originally also a proper name. Sometimes, names of particular persons or places may also become common words, as illustrated by the word *sandwich*, originating from an Earl of Sandwich, who is claimed to have been the first to put a slice of meat between two slices of bread.

1.4.2 Semantic change

Semantic change – a change in the meaning of words – is often difficult to distinguish from lexical change: the word *sandwich*, discussed above, is a good example. Generally, however, we talk about semantic change when a word acquires a new meaning, rather than when a new word is born. After all, the common noun *sandwich* is clearly a distinct word from the proper name *Sandwich*, from which it derives. Semantic change can be described as belonging to one of the following types.

1. **Semantic shift** happens when a word's meaning simply becomes different. The word *gay*, for example, used to mean 'merry, happy', but nowadays, it is generally used in the sense of 'homosexual'. The word *silly*, to take another example, used to mean 'blessed', then it came to mean 'clean, simple', and finally, 'stupid'. In Hungarian, the word *holott* used to mean 'where' (as a relative pronoun, corresponding to present-day Hungarian *ahol*). Semantic shift, however, can be of special types, discussed below.

2. **Broadening of meaning** takes place when a word acquires a more general meaning. In Old English, for example, the noun *dog* meant a special type of dog; now it refers to any animal belonging to the species.

3. **Narrowing of meaning** is just the opposite: a word with a more general meaning comes to refer to something more specific. In Early Modern English, for example, the noun *meat* meant 'food', whereas now it has come to mean a particular type of food.

4. **Amelioriation** is the term used when a word "rises" in its linguistic or social status. In Old English, for example, *queen* simply meant 'woman', then 'noble woman', and now, it means 'female monarch'. Similarly, *lord* used to mean (in Old English) 'leader of the household', whereas it now refers to a person of a high social rank. In Hungarian, *nagyszerű*, now meaning 'wonderful', used to mean simply 'large, extensive'.

5. **Deterioration** is the opposite process. It can be illustrated by the word *worm*, for instance, which used to mean 'dragon', but now has a much less highly regarded meaning. Hungarian *asszony* 'married woman', a loan from Persian, originally meant 'queen, princess' – the development is the opposite of what happened to the English word *queen*!

Needless to say, sometimes it is difficult to place a particular semantic change in a given category. The example given above, namely the word *silly*, may, after all, be regarded as an instance of deterioration, too.

1.4.3 Morphological change

Changes in the morphological system of a language can be quite complex, and they will be illustrated in detail in later chapters. In what follows, I provide a brief outline. Also, it must be noted that morphological change is often connected to, and difficult to separate from, syntactic change.

1. The most important type of morphological change is **analogy**, which basically means that a word assumes a new form based on a majority pattern. For example, the past and past participle forms of *help* used to be *holp* and *holpen*, respectively. On the analogy of regular verbs, however, the form *helped* was created (during the Early Modern English period), by now completely replacing the old forms. The appearance of regular past tense forms such as *dreamed* or *spelled* alongside older *dreamt* and *spelt* – to mention current examples - are other instances where a regular pattern exerts influence on irregular items. It would be mistaken, however, to think of analogy as simply "regularization". Sometimes, the opposite may happen, too, as in the case of the verb *catch*, which was originally a regular verb, but on the analogy of similarly sounding verbs such as *teach* it acquired the past (and past participle) form *caught*.

2. Changes in the inflectional or derivational morphology of a language can take place in other ways, too. One instance is the loss of certain morphological categories. In Early Modern English, for example, verbs – corresponding to personal pronouns, mentioned above – distinguished 2nd person singular and plural: in the singular, the suffix *-(e)st* was generally used in the indicative mood, so *thou comest* (singular, cf. Hungraian *te jössz*) was different from *you/ye come* (plural, cf. Hungarian *ti jöttök*). The picture is quite complex, though, and it will be discussed in detail in the next chapter; the point is that the singular vs. plural distinction has been lost in second person personal pronouns and verbs. On the other hand, new distinctions may arise, too: for instance, the pronoun *it* used to have a possessive form *his* – identical to the possessive form of *he*. The form *its*, found in present-day English, arose during the Early Modern English period, gradually replacing *his*.

An example in the change of the derivational morphology of English is the gain of the suffix *–able*: originally, it was found in loanwords from French and Latin, as in *changeable, countable,* etc., but it has become a productive derivational suffix by now, freely added to any transitive verb, including ones of native English origin, as illustrated by words such as *readable, eatable,* and so on. Conversely, the suffix *–th,* as in *width, length,* etc., used to form nouns from adjectives (cf. *wide* and *long*), which used to be relatively productive in the very early history of English, has been lost as a productive suffix, surviving only in a small number of relic forms. A similar example from Hungarian is the Locative (= indicating place) suffix *-t/-tt*, nowadays only found with some city names (e.g. *Pécsett, Kolozsvárt*) as well as some other relic forms such as *helyett* 'instead of'.

3. It often happens that something that used to be an independent word comes to be constantly attached to another word of a given category, losing its independent status and becoming an affix. The suffix *–less,* as in *careless,* was itself originally an adjective meaning 'free of something', so *careless* meant 'free of care'. In time, the adjective itself was lost, but it has survived as a productive suffix (cf. also *headless, countless,* etc.). This type of process is often termed **morphologization**, referring to the fact that a lexical content word becomes a bound morpheme, i.e. an affix.

1.4.4 Syntactic change

As I have mentioned, syntactic change is often difficult to separate from morphological change. A change in either system may result in changes in the other. In Old English, for example, there used to be a Dative case marked by a suffix added to nouns (and adjectives, pronouns, and determiners). So, the noun phrase *se wisa wer* 'the wise man' had the Dative form *ðæm wisan were* 'to the wise man'. In Old English then, a special case form was used, which, however, was lost, and, as the example shows, this loss has resulted in the more widespread use of the preposition *to* in order to take over the function of the lost Dative case. On the other hand, the more extensive use of the preposition *to* may have contributed to the loss of the Dative, making its use superfluous in many cases.

The syntax of a language is so complex that it would be a hopeless enterprise to give an overview of all aspects of syntactic change. I provide two examples only; further instances will be found in later chapters.

1. A phenomenon related to morphologization (discussed in the previous section) is **grammaticalization**, a process whereby a lexical content word loses its lexical meaning and becomes a grammatical function word. An example is provided by English *will,* originally a full verb meaning 'to want', so that a phrase like *I will go* meant 'I want to go'. In time, however, *will* has become an auxiliary denoting future tense, that is, a grammatical category.

Similarly, the negating particle *not* derives from *nought*, meaning 'nothing', so that *I will not* originally meant 'I want nothing'!

2. Spectacular instances of syntactic change can be found in the case of word order. In present-day English, the general word order of declarative clauses is SVO, i.e. Subject – Verb – Object, as in *The boy found the dog*. The same word order is found in subordinate clauses, too, witness *I know that the boy found the dog*. In subordinate clauses, however, there used to be another possibility in Old English: SOV, i.e. Subject – Object – Verb, so the previous example might as well have been *I know that the boy the dog found* – a word order which would clearly be impossible in Modern English.

1.4.5 Phonological change

This is perhaps the best studied aspect of language change, being essential for the historical linguist. The most important point to emphasize is that phonological change tends to be **regular**: this means that it is sounds and sound patterns that change, and more rarely do individual words change their pronunciation. In modern standard British RP, for example, Tense vowels fall into two groups, Plain and Broken (for the distinction, see, for example, Nádasdy 1996). The vowel **iː**, as in *feed*, is Plain Tense, for example, while the word *beer* illustrates it Broken Tense counterpart, i.e. **ɪə**. In Early Modern English, both words had an **iː**, furthermore, *r* was still pronounced in all positions, unlike in modern RP. Schematically, the differences can be represented as follows:

(2)	*feed*	*beer*
Early Modern English	**fiːd**	**biːr**
Modern RP	**fiːd**	**bɪə**

During the 18th century, however, a change known as Breaking took place, affecting all Tense vowels (not only **iː**, but I neglect the other ones here; the issue will be discussed in Chapter 3 in detail). Breaking, to put it simply, means that the Plain Tense vowel, that is, a long monophthong in this case, becomes a diphthong ending in a schwa, so **iː** > **ɪə**. The point is that this process takes place before a following **r**, that is why it does not happen in *feed*. The regularity of sound change means that Breaking *always* happens before **r**, and, conversely, it *never* takes place if the vowel is *not* followed by **r**. There is no exception to this rule, that is, it is a regular sound change. Sound change, therefore, takes place according to strict phonetic conditioning, rather than affecting words in a random fashion. Further examples include *bee, beat, reason*, with a Plain Tense **iː**, versus *beard, here,weary* with a broken **ɪə**.

It is to be noted that after Breaking had taken place, another change happened: **r** was dropped before a consonant or at the end of the word (but not before a vowel), so *beer, here, beard*, etc. no longer have a pronounced **r** (although word-final **r** is still pronounced if the following word begins with a vowel, as in *beer is* or *here I am*).

A sound may also change in all positions: in this case, we talk about **unconditioned** change. Such changes are also regular, since they take place in all words where the given sound is found. In the early 17th century, for example there was a diphthong **əɪ** in English, found in words like *time, ride, my*, etc. By the 18th century, it became **aɪ** in all cases, i.e. in every word where it was found. There is no exception to this rule.

Sometimes, however, a change in the pronunciation of words may be irregular, affecting certain words but not others, even though the phonetic environment is the same. In

Middle English, word such as *dead, head, bread* were prounounced with a long vowel (= **ɛː**). During the Early Modern English period, however, the vowel was shortened, so these words are now pronounced with a short **e** sound. Common to these words is the fact that the long vowel is followed by **d**, so it may be concluded that shortening took place before this consonant. Yet, there are many words where the long vowel is followed by **d**, too, but no shortening happens, such as *read, bead, lead* (verb), etc. (In fact, shortening could take place before some other consonants, too, such as **θ**, as in *breath, death*, etc., but there again, it is not consistent, cf. *heath*, with a long vowel.) May I note that Middle English **ɛː**, when not shortened, regularly becomes Modern English **iː**; this issue will be discussed later on.

The reasons why some sound changes are not regular is a complex and hotly debated issue among historical linguists. We will not deal with it here, but see the suggested readings at the end of the chapter.

Sound change, however, may not simply result in a difference in pronunciation: it may also affect the phonological system of the languague. Let us now illustrate this by two examples: **merger** and **split**.

1. **Merger** is said to take place when a phoneme changes into another one – one that already exists, that is, two phonemes fall together. In Early Modern English (more precisely, in the early 17th century), words such as *meat, peace, sea*, etc. were pronounced with a long vowel **eː**, while words like *meet, piece, see* had an **iː** (just like today). During the 17th century, however, the vowel **eː** became **iː**, while old **iː** remained unaltered. As a result, the two vowels fell together, and pairs such as *meet – meat, piece – peace, see – sea* are now homophones, while they were minimal pairs in Shakespeare's time (so, for example, *sea* and *see* did not rhyme, while they do so today). In other words, the result of merger is that one phoneme is completely lost, i.e. the number of phonemes decreases.

An example from Hungarian may also be illuminating. In early Hungarian, there was a sound dentoted by the phonetic symbol ʎ, a palatal ("soft") **l**, spelt *ly*, which has become **j**, merging with the already existing **j**, so the words *szablya* and *szabja*, to give an example, are now pronounced the same (note that the spelling still reflects the earlier difference).

2. **Split** is the opposite of merger: it means that a phoneme develops into two (or sometimes more, but I stick to two for simplicity's sake). Remember the change known as Breaking, discussed above: Tense vowels are broken before a following **r**, but remain Plain Tense otherwise. While Breaking produces new sounds, it does not produce new phonemes, for the Broken vowels are simply allophones of the Plain ones, the former appearing before **r** – that is, for example, **iː** and **ɪə** are in complementary distribution. This is illustrated in line 1. in table (3) below: although *bee* and *beer* have different vowels, the difference is the automatic consequence of the absence vs. presence of a following **r**, so the two words are not minimal pairs. The same goes for *bead* and *beard*. After Breaking, however, R-dropping takes place, so the **r** is lost in *beer* and *beard*. This is shown in line 2. of the same table. The result is that *bee* and *beer* (just like *bead* and *beard*), are now a minimal pair, differing in the vowel only. By defintion, two sounds are different phonemes if they distinguish words, so the Plain Tense and the Broken Tense vowels are now different phonemes, rather than allophones of one phoneme: the original single phoneme has split into two, and a new phoneme has been born.

(3)

Early 18th c. situation:	*bee*	*bead*	*beer*	*beard*
	biː	**biːd**	**biːr**	**biːrd**
1. Breaking (**iː** > **ɪə/___r**)	**biː**	**biːd**	**bɪər**	**bɪərd**
2. R-Dropping	**biː**	**biːd**	**bɪə**	**bɪəd**

This concludes our discussion of the essentials of linguistic change. Of course, there is much more to say, and I have indeed tried to give an overview, rather than a complete discussion, often simplifying and omitting things. Please check the suggested readings below.

Suggested reading

In this section, as well as the *Suggested reading* section at the end of every other chapter, the author(s) and the year of publication are given; please check them in the Bibliography at the end of the book. Another note: the entries of Wikipedia on the Internet contain a huge amount of useful information, but they must be handled with caution, and the info you find on Wikipedia is best checked in other sources!

The literature on language change is enormous. For beginners, perhaps the best and most accessible one is Aitchison (2001). Other works – also intended for beginners, but assuming a more-than-basic acquaintance with linguistics – include McMahon (1994), Bynon (1983), Lehmann (1993) and Campbell (2004).

For general introductions to the history of English, see Part I of the Bibliography.

2 Early Modern English Grammar

2.1 General notes

As mentioned in the previous chapter, Early Modern English is taken here to be the period between 1500 and 1800. Needless to say, during three centuries, the language changed quite a lot, so the English of the 16th century is not the same as that of the 18th. In this chapter, when I talk about Early Modern English (henceforth abbreviated to EModE), I refer to the language of the late 16th – early 17th centuries, i.e. the time of Shakespeare (the best known author of the period).

To illustrate EModE, I will use two texts. The first of these is a chapter from the Gospel according to Matthew (Chapter 5, to be precise, part of what is known as the Mountain Sermon), from the 1611 translation of the Bible known as the King James version (also as Authorized version). This translation is still the classical one used in English-speaking countries, although, of course, more modern translations are also available. It represents a somewhat old-fashioned language relative to its time. This is understandable: sacred texts tend to be conservative linguistically speaking. Just think of the Lord's Prayer in Hungarian, in which – up to the present – some archaic features are found (such as *szenteltessék meg a te neved* 'may your name be made holy'; *szenteltessék* is an archaic passive form no longer found in the living language). Roughly, the language of the King James version reflects the living usage of the early 16th century, though dating from the early 17th. The second text I will use is a passage from Shakespeare's *Julius Caesar*, usually dated to 1599. Though somewhat earlier than the King James version of the Bible, linguistically, it is more "up-to-date". Needless to say, the differences between the two texts are relatively minor, since language does not change too much in a century's time, but some of these differences are of particular interest.

I give these texts using a modernized spelling, although the spelling conventions of EModE were a bit different from those today. The reason for this is that EModE texts, whether the Bible or the works of Shakespeare (or whatever else), are usually printed according to present-day spelling rules, so that you are unlikely (unless specifically intending to study the original texts) to read them according to the way they were spelt at the time. Nonetheless, some spelling differences will be noted during our discussion of EModE. In terms of punctuation or the use of capital vs. lowercase letters, however, even the modernized texts may not follow the usual modern conventions.

Please read the passage from the King James Bible first. Do not worry if you find it difficult – it is not very easy for native speakers of English, either. At the end of the chapter, you will find, in the form of an appendix, a recent translation, which will be of help. Also, if your native language is other than English (and that is likely to be the case for most students reading this book), you may also check a translation in your own language.

After the passage, you will find a description of the most important features of EModE grammar, where examples from the passage will be used whenever possible. (If an important feature of EModE cannot be illustrated by the text, I will give examples taken from elsewhere.)

Finally, the Shakespearean text mentioned above will be presented, and I will point out the chief differences between the language of the King James version and the language of Shakespeare.

2.2 Matthew 5 from the King James version of 1611

1: And seeing the multitudes, he went up into a mountain: and when he was set, his disciples came unto him:
2: And he opened his mouth, and taught them, saying,
3: Blessed are the poor in spirit: for theirs is the kingdom of heaven.
4: Blessed are they that mourn: for they shall be comforted.
5: Blessed are the meek: for they shall inherit the earth.
6: Blessed are they which do hunger and thirst after righteousness: for they shall be filled.
7: Blessed are the merciful: for they shall obtain mercy.
8: Blessed are the pure in heart: for they shall see God.
9: Blessed are the peacemakers: for they shall be called the children of God.
10: Blessed are they which are persecuted for righteousness' sake: for theirs is the kingdom of heaven.
11: Blessed are ye, when men shall revile you, and persecute you, and shall say all manner of evil against you falsely, for my sake.
12: Rejoice, and be exceeding glad: for great is your reward in heaven: for so persecuted they the prophets which were before you.
13: Ye are the salt of the earth: but if the salt have lost his savour, wherewith shall it be salted? it is thenceforth good for nothing, but to be cast out, and to be trodden under foot of men.
14: Ye are the light of the world. A city that is set on an hill cannot be hid.
15: Neither do men light a candle, and put it under a bushel, but on a candlestick; and it giveth light unto all that are in the house.
16: Let your light so shine before men, that they may see your good works, and glorify your Father which is in heaven.
17: Think not that I am come to destroy the law, or the prophets: I am not come to destroy, but to fulfil.
18: For verily I say unto you, Till heaven and earth pass, one jot or one tittle shall in no wise pass from the law, till all be fulfilled.
19: Whosoever therefore shall break one of these least commandments, and shall teach men so, he shall be called the least in the kingdom of heaven: but whosoever shall do and teach them, the same shall be called great in the kingdom of heaven.
20: For I say unto you, That except your righteousness shall exceed the righteousness of the scribes and Pharisees, ye shall in no case enter into the kingdom of heaven.
21: Ye have heard that it was said by them of old time, Thou shalt not kill; and whosoever shall kill shall be in danger of the judgment:
22: But I say unto you, That whosoever is angry with his brother without a cause shall be in danger of the judgment: and whosoever shall say to his brother, Raca, shall be in danger of the council: but whosoever shall say, Thou fool, shall be in danger of hell fire.
23: Therefore if thou bring thy gift to the altar, and there rememberest that thy brother hath ought against thee;
24: Leave there thy gift before the altar, and go thy way; first be reconciled to thy brother, and then come and offer thy gift.
25: Agree with thine adversary quickly, whiles thou art in the way with him; lest at any time the adversary deliver thee to the judge, and the judge deliver thee to the officer, and thou be cast into prison.
26: Verily I say unto thee, Thou shalt by no means come out thence, till thou hast paid the uttermost farthing.
27: Ye have heard that it was said by them of old time, Thou shalt not commit adultery:

28: But I say unto you, That whosoever looketh on a woman to lust after her hath committed adultery with her already in his heart.
29: And if thy right eye offend thee, pluck it out, and cast it from thee: for it is profitable for thee that one of thy members should perish, and not that thy whole body should be cast into hell.
30: And if thy right hand offend thee, cut if off, and cast it from thee: for it is profitable for thee that one of thy members should perish, and not that thy whole body should be cast into hell.
31: It hath been said, Whosoever shall put away his wife, let him give her a writing of divorcement:
32: But I say unto you, That whosoever shall put away his wife, saving for the cause of fornication, causeth her to commit adultery: and whosoever shall marry her that is divorced committeth adultery.
33: Again, ye have heard that it hath been said by them of old time, Thou shalt not forswear thyself, but shalt perform unto the Lord thine oaths:
34: But I say unto you, Swear not at all; neither by heaven; for it is God's throne:
35: Nor by the earth; for it is his footstool: neither by Jerusalem; for it is the city of the great King.
36: Neither shalt thou swear by thy head, because thou canst not make one hair white or black.
37: But let your communication be, Yea, yea; Nay, nay: for whatsoever is more than these cometh of evil.
38: Ye have heard that it hath been said, An eye for an eye, and a tooth for a tooth:
39: But I say unto you, That ye resist not evil: but whosoever shall smite thee on thy right cheek, turn to him the other also.
40: And if any man will sue thee at the law, and take away thy coat, let him have thy cloak also.
41: And whosoever shall compel thee to go a mile, go with him twain.
42: Give to him that asketh thee, and from him that would borrow of thee turn not thou away.
43: Ye have heard that it hath been said, Thou shalt love thy neighbour, and hate thine enemy.
44: But I say unto you, Love your enemies, bless them that curse you, do good to them that hate you, and pray for them which despitefully use you, and persecute you;
45: That ye may be the children of your Father which is in heaven: for he maketh his sun to rise on the evil and on the good, and sendeth rain on the just and on the unjust.
46: For if ye love them which love you, what reward have ye? do not even the publicans the same?
47: And if ye salute your brethren only, what do ye more than others? do not even the publicans so?
48: Be ye therefore perfect, even as your Father which is in heaven is perfect.

2.3 Early Modern English grammar

In this section, we discuss the chief grammatical properties of EModE inasmuch as they are different from standard varieties of Present-day English (hence abbreviated PdE).

2.3.1 Morphology

The most important morphological differences between EModE and PdE involve (1) pronouns, (2) verbs. I start the discussion with these, therefore.

2.3.1.1 Personal pronouns

The personal pronouns of EModE are seemingly the same as today, but there are some important differences – and appearances might be deceptive. Today, the system of personal pronouns looks as shown in table (4) below. Notes: Sg = Singular, Pl = Plural; numbers indicate person (first, second, and third), while cases are: Nom = Nominative, Acc = Accusative, Gen = Genitive.

(4)

	SG1	SG2	SG3			PL1	PL2	PL3
NOM	*I*	*you*	*he*	*she*	*it*	*we*	*you*	*they*
ACC	*me*	*you*	*him*	*her*	*it*	*us*	*you*	*them*
GEN	*my/mine*	*your(s)*	*his*	*her(s)*	*its*	*our(s)*	*your(s)*	*their(s)*

As you can see, the 2nd person personal pronouns are identical in the singular and the plural: *you* can refer to one person or more. These are shown in the shaded boxes. As mentioned in Chapter 1, this is an important difference between EModE and PdE, but there are other, minor differences, too. Let us now see a similar table, showing EModE personal pronouns:

(5)

	SG1	SG2	SG3			PL1	PL2	PL3
NOM	*I*	*thou*	*he*	*she*	*it*	*we*	***ye***	*they*
ACC	*me*	*thee*	*him*	*her*	*it*	*us*	*you*	*them*
GEN	*my/mine*	*thy/thine*	*his*	*her(s)*	***his***	*our(s)*	*your(s)*	*their(s)*

The singular 2nd person pronouns are clearly different from the plural 2nd person ones. Furthermore, the boldfaced items (the Genitive form *his* of *it* and the Nominative *ye* of *you*) are also different. Let us now look at these differences in detail.

First, *thou* (and its various forms) diifer from *ye* (and its various forms) in that *thou* is singular, while *ye* is plural. To compare the situation to Hungarian, *thou* corresponds to Hungarian *te*, while *ye* corresponds to *ti*. Look at the following examples from the King James text:

(6) *Ye are the salt of the earth* (verse 13) 'You are the salt of the earth'
(7) *Thou shalt not kill* (verse 21) 'You shall not kill', i.e. 'Do not kill'

In (6), several persons are addressed, while in (7), only one; in PdE, *you* would be used in both cases. (Note: *shalt* is a special form of *shall*, used when the subject is *thou*; more on this below.) In Hungarian, you find a similar situation - *Ti vagytok a föld sója* vs. *Ne ölj!* – in the former, the subject is plural, while in the second, it is singular. Another, quite famous example is provided by Shakespeare, from *Romeo and Juliet*, when Juliet asks,

(8) *Oh Romeo, wherefore art thou Romeo?*

This can be translated into PdE as 'Oh Romeo, why are you Romeo' – or as 'Ó Rómeó, miért vagy te Rómeó' into Hungarian. (The form *art*, just like *shalt*, is the special 2nd person singular form of *be* in the Present Indicative, corresponding to Hungarian *vagy*, and it has nothing to do with the noun *art* 'művészet' at all.) In sum, EModE makes a distinction between singular and plural in the second person, a distinction no longer found in English.

However, the picture is more complex, and take heart now: *ye* (and its various forms) can indeed be used as a singular form in EModE. No example of this is found in the King James version[2], but look at the following example, taken from Shakespeare's *Love's labour's lost*:

(9) *I praise God for <u>you</u> sir, <u>your</u> reasons at dinner have been sharp*
'I praise God for you, sir; your remarks at dinner were sharp'

The underlined forms clearly refer to one person, that is, they have a singular reference: this is clear from the word *sir*, a singular form. The speaker, then, addresses only one person. Yet, *you* and *your* are plural forms; why are they used, then? The reason for this is that 2nd person plural pronouns could also be used with a singular reference if you wanted to address someone politely, i.e. you wanted to use **formal address** (= *magázás* in Hungarian). *Ye* (and its various forms) were opposed to *thou* (and its forms) in two ways:

1. *Thou* was singular only, while *ye* could be either singular or plural, but
2. *ye* was only used as a singular only to express formal address.

To present a Hungarian parallel, we can say that *thou* corresponds to *te*, while *ye* corresponds to (1) *ti*, (2) *ön/maga, önök/maguk*. Note that EModE makes no distinction in number in formal address: *ye*, when used as a polite form, can be either singular or plural (in other words, EModE does not distinguish what corresponds to Hungarian *ön* vs. *önök*.

Similarly to Hungarian, the formal address was used when talking to someone's superiors (such as a servant to his master) or some stranger you did not want to offend, etc. The **informal address** (= Hungarian *tegezés*) was used when talking to a close friend, or when a master spoke to his servant, or when addressing children or other persons you were in intimate contact with (so Juliet uses *thou*, the informal form, to address Romeo, as shown by the example in (8) above). The forms of *thou* are gradually replaced by 2nd person plural forms, however, because formal address was used more and more widely: simply, people used it because it sounded more polite or elegant. By the 18th century, *thou* had been virtually lost, although it survives much longer in poetic usage, for example.

[2] This is because in Ancient Greek, the original language of the New Testament, the relevant distinction between formal and informal address (to be discussed right away) did not exist (neither did it exist in Latin, for that matter).

So finally, here is a completed table of EModE personal pronouns (the highlighting has now been omitted):

(10)

	SG1	SG2		SG3			PL1	PL2	PL3
		Informal	Formal						
NOM	*I*	*thou*	*ye*	*he*	*she*	*it*	*we*	*ye*	*they*
ACC	*me*	*thee*	*you*	*him*	*her*	*it*	*us*	*you*	*them*
GEN	*my/mine*	*thy/thine*	*your(s)*	*his*	*her(s)*	*his*	*our(s)*	*your(s)*	*their(s)*

A further point to discuss is the difference between Nominative and Accusative forms, like between *ye* and *you*, or *thou* and *thee*. This is parallel to the distinction between *I* and *me* or *she* and *her* (etc.) in PdE. Nominative forms are used when the pronoun is the subject, while Accusative forms are used when it is a (direct or indirect) object, or when part of a prepositional phrase (prepositions govern the Accusative case in EModE as well as in PdE). So, for example, you use *I* vs. *me* in PdE according to the following examples:

(11)
(a) *I like reading*. (= *I* is the subject)
(b) *She loves me*. (= *me* is the object)
(c) *My son takes after me*. (= *me* is preceded by a preposition)

The difference between *thou* vs. *thee* and *ye* and *you* is identical in EModE (examples from the King James text):

(12)
(a) *Thou shalt not kill* (verse 21) = subject
(b) *lest at any time the adversary deliver thee to the judge* (verse 25) = object
(c) *Verily I say unto thee* 'I truly say to you' (verse 26) = after the preposition *unto* 'to'

(13)
(a) *Ye are the salt of the earth* (verse 13) = subject
(b) *when men shall revile you* (verse 11) 'when people insult you' = object
(c) *the prophets which were before you* (verse 12) = after the preposition *before*

It must be noted that – in the case of *ye* and *you* – this distinction is an archaic feature of the text: by Shakespeare's time, the two forms become interchangeable, and *ye* becomes more rare, replaced by *you*. Think of the title of the play *Twelfth night; or what you will* (Hungarian *Vízkereszt, vagy amit akartok*, where *you* is used as a subject. We will see further examples in the Shakespearean text below.

Let us now turn to other differences between EModE and PdE. First, observe the forms of the pronoun *it*: you will find that its Genitive form is *his*, rather than *its*, as in PdE. Look at the following example from the King James text:

(14) *if the salt have lost his savour, wherewith shall it be salted?* (verse 13)
'if the salt has lost its taste, what will it be salted with?'

The form *its* came into existence during the EModE period, not yet used in this text, but already existing is popular speech. It is interesting to note that it is variably spelt *its* or *it's*, but later on, the form *its* became the norm, *it's* being reserved for the contracted form of *it is* (or *it has*). Another point to note in connection with the pronoun *it* is that it is often abbreviated to *'t*, as in *'twil* 'it will'.

Second, the difference between *my* vs. *mine*, as well as *thy* vs. *thine*, requires specific mention. As *thy/thine* no longer exist in PdE, I will first illustrate the difference between *my* and *mine*. In PdE, the choice between the two is grammatically determined: *my* is used in an attributive function (i.e. as a determiner, preceding the modified noun), as in *my car*, *my apple*. *Mine*, however, is used in a predicative function, as in *This car is mine.* In EModE, *mine* was also used in the same function, but it could also be used attributively, notably, when the noun began with a vowel. This is like the choice between *a* and *an* in PdE (and in EModE, too): you say *a car* but *an apple*, because *apple* begins with a vowel. Similarly, in EModE, you said *my car* – but *mine apple*. The same goes for *thy/thine*; witness the following example from the King James passage:

(15)
(a) *Leave there <u>thy gift</u> before the altar* (verse 24)
(b) *Agree with <u>thine adversary</u> quickly* (verse 25)

In (15a), *thy* is used because the noun it modifies (= *gift*) begins with a consonant, but in (15b), *thine* is found because *adversary* begins with a vowel.

It must be noted that in EModE, words beginning with **h** behaved as if they actually began with a vowel. Look at the example below:

(16) *A city that is set on <u>an hill</u> cannot be hid* (verse 14)
'A city that is placed/built on a hill cannot be hidden'

Today, we would say *a hill*, but in EModE, *an* was used – just like in *an apple*. The same is true for the use of *mine* and *thine*, so you would say, in EModE, *mine hill* or *thine hill*. Nonetheless, the use of *my* and *thy* was becoming more and more widespread in EModE in an attributive function. Examples illustrating this are *my own* and *my head*, for expected *mine own* and *mine head*, found in some of Shakespeare's works.

A final note: the pronoun *them* is often found as *'em*. This form originates from older *hem*. Indeed, it is still widespread in non-standard varieties of English, but no longer existing in standard English.

2.3.1.2 Demonstrative pronouns

Demonstrative pronouns in PdE are *this/these, that/those*. In EModE, there was another demonstrative, appearing in various forms: *yon, yond, yonder*, meaning 'that over there', suggesting a distance from both the speaker and the hearer.

2.3.1.3 Verbs

Again, the morphology of verbs in EModE is quite similar to PdE, but there are some important differences, too. The following table shows a typical verbal paradigm in EModE:

(17) The Present and Past forms of the verb *love* in EModE

	Indicative mood	**Subjunctive mood**
Present tense		
1st person singular	*I love*	*I love*
2nd person singular	*thou lovest*	*thou love*
3rd person singular	*he/she/it loveth*	*he/she/it love*
Plural (all persons)	*we/ye/they love*	*we/ye/they love*
Past tense		
1st person singular	*I loved*	*I loved*
2nd person singular	*thou lovedest*	*thou loved*
3rd person singular	*he/she/it loved*	*he/she/it loved*
Plural (all persons)	*we/ye/they loved*	*we/ye/they loved*

The shaded boxes show forms which are no longer found in English. In the 2nd person singular forms, this is understandable: as dicussed above, this category has been lost (being replaced by the 2nd person plural), so the corresponding verb forms have disappeared, too. I must also note that in the Subjunctive, there is no distinction in person/number. As for the use of the Subjunctive mood, see below (the section discussing syntax).

Two important points are to be noted.

First, verbs generally take the suffix *–est* (often *–st*) in the Indicative mood, when the subject of the clause is the pronoun *thou*. This is illustrated by the example in (18):

(18) *<u>thou canst</u> not make one hair white or black* (verse 36)
'you cannot make one hair white or black'

As this example shows, the verb *can* takes the suffix *–st*, because the subject is the 2nd person singular pronoun *thou*.

Some verbs, on the other hand, take the suffix *–t* in the 2nd person singular, like *shalt, art, wilt*; they represent a minority pattern, however, and they are generally auxiliaries. (See the examples in (7) and (8) above.)

Second, note that in the 3rd person singular of the Present Indicative, the ending *–eth* (or *–th*) is used instead of PdE *–s* or *–es*. So, EModE *giveth* corresponds to PdE *gives*, or *hath* corresponds to PdE *has*. Look at the examples in (19):

(19)
(a) *<u>it giveth</u> light unto all that are in the house* (verse 15)
(b) *<u>whosoever looketh</u> on a woman to lust after her <u>hath</u> committed adultery with her already in his heart* (verse 28)
'anyone who looks at a woman lustfully (= lusting for her) has committed adultery with her already in his heart'

The underlined parts contain a 3rd person singular subject, and, accordingly, the suffix *-(e)th* is used. The suffix *-(e)s*, as in PdE *gives, has, looks*, etc., originates from the North of England, where it had been used as early as in Old English times, gradually spreading towards the South. It is not yet found it the text of the King James version, but it is very frequent in Shakespeare's works (see the passage from Julius Caesar below), indicating that it was becoming more and more widespread in London – and the South of England in general – at around 1600. By the end of the 17th century, *-(e)s* completely replaced the old *-(e)th* in general

use, although it continued to be used in archaic, especially poetic, usage long after that (the prestige of the King James Bible may have contributed to this, to mention one factor).

As for the past tense forms of verbs, some of them had variants. The verb *write*, for instance, had the past tense form *writ* alongside *wrote*. (Only the latter survives by today, of course.) The reason for this variation goes back to very ancient times; originally, it reflects a difference in person and number, but in PdE, it is only the verb *be* which has preserved this distinction (cf. *I/he/she/it* <u>*was*</u> but *you/we/they* <u>*were*</u>).

The Present Participle of verbs (e.g. *reading, being, writing*) was formed in the same way as today, although there was a difference in use, illustrated by the example below:

(20) *Rejoice, and be* <u>*exceeding*</u> *glad* (verse 12)

Today, we would use the adverbial form *exceedingly*, but in EModE, the Present Participle could be used in that function without the adverbial suffix *–ly*.

The Past Participle of most verbs was also identical to the PdE ones, although there are, of course, sporadic differences. Consider the example in (21):

(21) *A city that is set on an hill cannot be* <u>*hid*</u> (verse 14)

Today, the form *hidden* would be used. The original form is *hid*, but on the analogy of pairs such as *write – written* or *ride – ridden*, where the suffix *–en* is ancient, *hid* was replaced by *hidden*. (Cf. Section 1.4.3.) In some cases, irregular Past Participles have been replaced by regular ones (such as *holpen* by *helped*, both existing in EModE, but only the latter has survived); conversely, a small number of verbs have acquired an irregular Past Participle, such as *dig*, whose Past Participle is *dug* today (irregular), but was regular (= *digged*) in EModE.

Finally, a nice example is provided by the variant forms of the Past Participle of *be* in EModE, viz. *been* (pronounced **biːn**) and *bin* (pronounced **bɪn**). In PdE, the spelling is uniformly *been*, but the EModE pronounced variants still survive, **biːn** being more common in Britain, while **bɪn** being usual in America!

2.3.1.4 Other parts of speech

Nouns, adjectives, prepositions, etc. show no essential difference relative to PdE. Some of them, of course, have been lost, as shown by the case of *unto*, a preposition meaning *to* (as in *verily I say unto you* 'I truly say to you); you can find this preposition several times in the King James text. May I note that before an Infinitive (as in *to look*) only *to* was possible in EModE, too. Some irregular noun plurals have become regularized, e.g. *eyen* (> *eyes*), or *horse* (> *horses*; this word originally had a suffixless or **zero plural**, like PdE *sheep*).

A systematic difference between EModE and PdE concerns a matter of spelling. In PdE, the Genitive forms of nouns are written with an apostrophe, e.g. *stone's, boys'*, etc. In EModE, this is not yet observed, so you would have *stones, boys* (identical to the non-genitive plural forms, which sound the same as the Genitive forms, anyway, up to the present). Since, however, you are more likely to read EModE texts with a modernized spelling, you will rarely come across such spelt forms.

2.3.2 Early Modern English syntax

The syntax of EModE is relatively similar to PdE, just like its morphology, but of course, there are some differences – and sometimes, even what looks similar may mean something different. The King James text illustrates these points in abundance. Let us now look at the most important features of EModE syntax.

2.3.2.1 The use of the auxiliary *do* and related issues

One of the crucial differences between EModE and PdE is in the use of *do* as an auxiliary. In PdE, *do* (and all its forms, i.e. also *does, did*) functions as a "dummy" auxiliary in the Simple Present and the Simple Past, as shown below in (22) – compare these to the examples in (23), where there is some other auxiliary:

(22)		(23)
(a)	*I do not like milk.*	(a) *I should not eat that cake.*
(b)	*Do you read novels?*	(b) *Can you come to the party?*
(c)	*She <u>does</u> love him.*	(c) *I <u>can</u> play the piano.*
(d)	*Little did he know about it.*	(d) *Little could we achieve.*

In (22a), *do* is used to express negation: it is a dummy auxiliary because it does not add anything to the meaning of the sentence: it has to be there because main verbs (such as *like*) cannot be negated directly, so you cannot say **I like not milk.* As shown by (23a), however, auxiliaries are directly negated by *not*: *I should not*, rather than **I do not should*. In (22b), *do* is used in a question, because main verbs cannot undergo Subject-Auxiliary conversion, so **Read you novels?* is ungrammatical; compare this to (23b), where *can*, an auxiliary, does indeed do so (cf. the impossible form **Do you can come to the party?*). In (22c), *does* is used to express emphasis, and accordingly, it is strongly stressed (shown by *does* being underlined). In the case of other auxiliaries, which are not dummy ones, so they must be present anyway, it is their being stressed only that betrays emphasis, cf. (23c): if *I can play the piano* is pronounced with an unstressed *can*, it is a neutral statement, but if *can* is stressed, the statement is emphatic (= *But of course I <u>can</u> play the piano*). In (22d) and (23d), we see an instance of an adverb (or adverbial phrase) being placed at the beginning of the sentence, after which Subject-Auxiliary conversion takes place. This happens quite rarely – other such adverbials include *hardly, no sooner, never, so, neither*, etc., - but the point is that if they are in sentence-initial position, you need Subject-Auxiliary conversion. In (23d), there is no problem, since *could* is an auxiliary so it can be converted; but in (22d), *did* is needed, because *know* is a main verb rather than an auxiliary, so it cannot undergo conversion (just like in (22b)): **Little knew he about it* would be impossible.

In EModE, the situation is rather different: *do* can be used – optionally – in such cases, i.e. in the Simple Present and the Simple Past, when there is no other auxiliary. Essentially, however, its use is truly optional. This is because in EModE, main verbs – unlike in PdE – can be directly negated by *not*, and can undergo conversion, too. Look at the following examples from the King James text:

(24)

(a) *<u>Think not</u> that I am come to destroy the law* (verse 17)
'Do not think that I have come to abolish the law'

(b) *<u>do not</u> even the publicans the same?* (verse 46)
'do not even the tax collectors do the same?'

(c) *so persecuted they the prophets which were before you* (verse 12)
'so did they persecute the prophets who were before you'

In these examples, there is no auxiliary at all: *think* is directly negated by a following *not* (24a); in (24b), *do* is not an auxiliary but a main verb (as shown by the translation), meaning 'perform'; in (24c), the verb *persecuted* undergoes conversion, i.e. it comes before the subject, because of *so* being placed at the beginning of the clause.

Of course, you can also find instances where *do* as a dummy auxiliary is used, see the following examples:

(25)
(a) *Blessed are they which do hunger and thirst after righteousness* (verse 6)
'blessed are those who hunger and thirst for righteousness'
(b) *Neither do men light a candle* (verse 15)

However, both sentences could be used without *do*, and the meaning would still be the same; (25b), this would result in a change in word order: the verb *light* would come before the subject, i.e. *Neither light men a candle* (cf. (24c) above). As I mentioned, in PdE, the use of *do* in declarative clauses expresses emphasis: *I love you* does not mean the same as *I do love you*. There is no such difference in EModE: the two mean exactly the same, *I do love you* expresses no emphasis at all.

2.3.2.2 Tense and aspect

In EModE, the tense/aspect system is basically the same as in PdE, with some differences, however.

1. The future tense is mostly expressed by *shall*, rather than *will*, in all persons. There are lots of examples in the King James text, cf. (26) below:

(26) *Blessed are the meek: for they shall inherit the earth.*
'Blessed are the meek: for they will inherit the earth.'

An interesting feature of EModE is that the future tense can be used in subclauses, such as in temporal clauses introduced by *when*. In PdE, the present tense is used in such cases, even if the verb refers to a future event. Look at the examples below:

(27) *when men shall revile you* (verse 11) 'when people insult you'

Shall, in fact, originally was a full verb expressing obligation – this use is still found in EModE, witness *Thou shalt not kill* 'You shall not kill'; indeed, it can still be used in the same sense in modern formal English, especially in legal texts (e.g. *the parties shall agree*).

The auxiliary *will*, however, is primarily still used as a full verb meaning 'to want', cf. *Twelfth night; or what you will* (Vízkereszt, vagy amit akartok). Sometimes, however, it is already found with a future meaning in Shakespeare's texts, but not yet in the King James Bible.

2. Perfect tenses are found in EModE, but their use is different: often a Present Perfect is used when we would use a Simple Past (or vice versa). In general, simple tenses are more frequently used, however. A Present perfect is sometimes found in the King James text, but it

is definitely rare. An example is provided by *it hath been said* in verse 31, or *ye have heard* in verse 27. The passage from Julius Caesar will contain further examples.

3. Progressive forms such as *I am coming* or *he has been doing* are indeed found (in fact, they appear already in Middle English), but they are rare, and simple forms are generally used instead.

All in all, the tense/aspect system of English as known today is still in a state of flux, and the system does not gain its present-day form until Later Modern English. This is shown by an interesting difference between American and British usage involving the use of the Present Perfect vs. the Simple Past: whereas a British speaker would say, for example, *Have you had lunch yet?*, Americans would prefer the Simple Past, i.e. *Did you have lunch yet?* Since in EModE, these two tenses were often used in free variation, one variant became the norm in America but the other one in Britain.

2.3.2.3 Mood

A marked difference between EModE and PdE is in the use of the Subjunctive. First, however, let's look at the forms of the subjunctive. Recall the paradigm in (17), repeated here as (28) for your convenience:

(28) The Present and Past forms of the verb *love* in EModE

	Indicative mood	**Subjunctive mood**
Present tense		
1st person singular	*I love*	*I love*
2nd person singular	*thou lovest*	*thou love*
3rd person singular	*he/she/it loveth*	*he/she/it love*
Plural (all persons)	*we/ye/they love*	*we/ye/they love*
Past tense		
1st person singular	*I loved*	*I loved*
2nd person singular	*thou lovedest*	*thou loved*
3rd person singular	*he/she/it loved*	*he/she/it loved*
Plural (all persons)	*we/ye/they loved*	*we/ye/they loved*

As you can see, the Subjunctive does not show any difference in person and number: in the present, it is basically the same as the Infinitive, i.e. *love* – this is true for all verbs, even for *be*, whose Present Subjunctive form is *be*. In the past, the same form is used throughout, too: this essentially means that the 2nd person singular lacks it characteristic *-(e)st* ending.

The chief difference between EModE and PdE is that the Subjunctive is used in *if* clauses, and sometimes in other subclauses, too. Look at the examples below:

(29)
(a) *if the salt <u>have</u> lost his savour* (verse 13) 'if the salt has lost its taste'
(b) *till all <u>be</u> fulfilled* (verse 18) 'until all is fulfilled'

As you can see, PdE would use the Indicative forms (= *has, is*) in these subordinate clauses (introduced by *if* and *till*, respectively), but EModE uses the Subjunctive. A particularly nice example is provided in (30) below:

(30) *Therefore <u>if thou bring</u> thy gift to the altar, and there [thou] <u>rememberest</u> that thy brother hath ought against thee* (verse 23)
'Therefore, if you bring your gift to the altar, and there you remember that your brother has anything against you'

In the first clause, introduced by if, the suffixless Subjunctive form is used, but in the second, which is not an *if* clause, you already have the 2nd persong singular form *rememberest*, i.e. an Indicative form. (The word *thou* is placed between brackets because it is omitted in the original text but is understood to be there as the subject.)

Finally, it has to be noted that the Past Subjunctive is used in EModE to express unreal conditions, where PdE uses the Indicative of the Simple Past, as in *<u>If I loved you</u>, I would marry you.* (But I do not love you.) The verb *be*, however, behaves specially: in all persons, the form *were* can be used in such conditional clauses, even when otherwise *was* would be used. So you can say *if I were you, if it were true*, etc. Of course, current usage tends to favour *was* (e.g. *if it was true*), especially in informal language, but the point is that this conditional use of *were* in all persons is a relic of the EModE situation: in EModE, *were* was the general Past Subjunctive form of be – recall that there was no person/number distinction in the Subjunctive!

2.3.2.4 Other

Let me now point out some other differences between EModE and PdE syntax.

1. The use of relative pronouns is a bit different. The relative pronoun *who* (as in PdE *The man who sat down beside me*) is relatively rarely used; on the other hand, *which* is commonly used to refer to human beings, as illustrated by the example in (31) below:

(31) *the prophets which were before you* (verse 12)
'the prophets who were before you'

Compare also the starting words of the Lord's Prayer as found in the King James version: *Our Father which art in heaven*, i.e. 'Our Father who are in heaven'.

2. Another feature characteristic of EModE is the use of the conjunction *that* after relative pronouns. No example is found in the King James passage, but an example from Shakespeare's *Julius Caesar*, to be presented in section 2.4. below, illustrates the point:

(32) *<u>When that</u> the poor have cried, Caesar hath wept*
'When the poor cried, Caesar wept'

In this example, you can find the relative pronoun *when* followed by *that*. In PdE, *that* is impossible in such cases. Nevertheless, the use of *that* in such cases is not obligatory, but only an option. If *when* (or other such pronouns such as *who*, *where*, etc.) are used as interrogative pronouns, rather than relative ones, *that* cannot be used, so **When that did you arrive?* would be ungrammatical. Let me also point out a difference between EModE and PdE, mentioned earlier, concerning the use of tenses: Shakespeare uses the Present Perfect, whereas today, as shown by the translation, we would use the Simple Past.

3. Personal pronouns are commonly used before a relative clause, where PdE would use a demonstrative pronoun. This is exemplified in (33):

(33) *Blessed are they which are persecuted for righteousness' sake* (verse 10)
'Blessed are those who are persecuted for righteousness' sake'

This example, again, also illustrates the previous point, i.e. the use of *which* to refer to humans. Sometimes, we would use some other construction, such as the pronoun *one*:

(34) *Give to him that asketh thee* (verse 42)
'Give to the one that asks you'

4. In PdE, the Perfect forms of verbs are always formed with the auxiliary *have*, cf. *I have eaten, I have come*. Compare, however, the example in (35):

(35) *Think not that I am come to destroy the law* (verse 17)
'Do not think that I have come to abolish the law'

In EModE, the auxiliary *be* is generally used with intransitive verbs (such as *come*); *have* is normal with transitive ones, which is easily understandable: *be* + a transitive verb would express a passive, so *I am eaten* does not mean the same as *I have eaten*. The use of *have*, however, is becoming more and more widespread during the EModE period. The use of *be* survives in some fixed expressions, such as *She is gone*, up to the present.

5. Some verbs have a different type of complement than today. Look at (36):

(36) *he maketh his sun to rise* (verse 45)
'he makes his sun rise'

Make, as a causative verb, requires a simple Infinitive today, i.e. one without a preceding *to*. Interestingly, *to* is still used in the passive, compare *I made him leave* vs. *He was made to leave.* Another example is provided in (37):

(37) *Blessed are they which do hunger and thirst after righteousness* (verse 6)
'blessed are those who hunger and thirst for righteousness'

The verbs *hunger* and *thirst* required the preposition *after* in EModE, while PdE would use *for*.

6. A special structure, frequent in EModE, is known as the **His Genitive**. No example is found in the King James text, but it occurs quite a lot elsewhere. It means that – instead of the Genitive ending *'s – his* is used, e.g. *the king his crown*, meaning 'the king's crown'. Sometimes, *her* and *there* are also used, e.g. *the queen her crown.*

There are, of course, other differences, too, but it would be useless to give a complete list of minor details; see the suggested readings at the end of the chapter for more exhaustive discussions.

2.3.3 Lexical differences

Changes in vocabulary are not as systematic as grammatical ones. Let me just give a couple of examples.

Some words have been lost or have become at least archaic. The word *ought*, meaning 'anything', is an example of a word that has been lost (it is not to be confused with the auxiliary *ought* as in *I ought to go*: the two have nothing to do with each other). Another example is the numeral *twain* (verse 41), an old form of *two*. The word *commandment*, too, has become archaic, being replaced by *command* – it survives, however, in the expression *The ten commandments* (Hungarian *tízparancsolat*). You can find further examples, too. It would be impossible, of course, to list all vocabulary changes here – you are advised to check a good dictionary, especially an etymological one, for words like these.

2.4 Another sample text: Antony's monologue from Julius Caesar by Shakespeare (Act 3, Scene 2)

Friends, Romans, countrymen, lend me your ears;
I come[1] to bury Caesar, not to praise him.
The evil that men do lives[2] after them;
The good is oft interred with their bones;
So let it be with Caesar. The noble Brutus
Hath told you Caesar was ambitious:
If it were so, it was a grievous fault,
And grievously hath Caesar answered it.
Here, under leave of Brutus and the rest -
For Brutus is an honourable man;
So are they all, all honourable men -
Come I to speak in Caesar's funeral.
He was my friend, faithful and just to me:
But Brutus says he was ambitious;
And Brutus is an honourable man.
He hath brought many captives home to Rome
Whose ransoms did the general coffers fill:
Did this in Caesar seem ambitious?
When that the poor have cried, Caesar hath wept:[3]
Ambition should be made of sterner stuff:
Yet Brutus says he was ambitious;
And Brutus is an honourable man.
You all did see[4] that on the Lupercal
I thrice presented[5] him a kingly crown,
Which he did thrice refuse:[6] was this ambition?
Yet Brutus says he was ambitious;
And, sure, he is an honourable man.
I speak not to disprove what Brutus spoke,
But here I am to speak what I do know.
You all did love him once, not without cause:
What cause withholds you then, to mourn for him?
O judgment! thou art fled to brutish beasts,
And men have lost their reason. Bear with me;

My heart is in the coffin there with Caesar,
And I must pause till it come back to me.

NOTES
The underlined parts, numbered, represent a selection as points of interest.

1. Note the use of the Simple Present here, while PdE would use either *I have come*, or, though less probably, *I am coming*; cf. section 2.3.2.2. above.

2. The suffix *-(e)s* is used in the 3rd person singular instead of *-(e)th*, not yet found in the King James text, though the old suffix is still found – look at the the form *hath*, occurring several times in this text. For the use of *-(e)s*, see also the form *says*, for older *saith*, occurring several times in this passage.

3. This example was explained in detail in section 2.3.2.4 above (see (32) and the explanations below).

4. Note the use of *you*, rather than *ye*, as a subject.

5. and 6. These show the optional use of *do* as an auxiliary: *presented* vs. *did refuse*. Recall that there is no difference in emphasis!

2.5 An interesting addition as a point of interest

This concludes our discussion of EModE grammar. Let me, however, quote a passage from John Keats' *Ode on a Grecian urn*, written in 1820. You will be able to recognize several EModE features in it, illustrating the fact that many EModE forms continued to be used in archaizing (chiefly poetic) usage for a long time after they had actually disappeared from the living language. Note: *dost* and *doth* are the Indicative Present 2nd and 3rd person forms of *do*, respectively.

(38) **The last verse of *Ode on a Grecian urn***
O Attic shape! Fair attitude! with brede
    Of marble men and maidens overwrought,
With forest branches and the trodden weed;
    Thou, silent form, dost tease us out of thought
As doth eternity. Cold Pastoral!
    When old age shall this generation waste,
        Thou shalt remain, in midst of other woe
Than ours, a friend to man, to whom thou say'st,
    "Beauty is truth, truth beauty" - that is all
        Ye know on earth, and all ye need to know.

It is interesting to point out that some EModE features are still occasionally used in PdE literature, sometimes in a mistaken way. In Jasper Fforde's novel entitled *Something rotten* (published in 2004), for example, you find the sentence *Puny […] agents, prepare to meet thy doom*, where the pronoun *thy* is used – from a historical point of view – quite erroneously: *agents* is a plural noun, so the expected pronoun form referring back to it would be *your* (= *prepare to meet your doom*). Whether the author uses *thy* because of his ignorance of this fact or delibaretely (for humorous purposes), I am not sure; at any rate, this example shows that for special effects, some EModE forms still persist, although they ceased to be a part of normal usage hundreds of years ago.

2.6 (APPENDIX) Matthew 5 from the New International Version of the Bible[3]

1 Now when Jesus saw the crowds, he went up on a mountainside and sat down. His disciples
came to him,
2 and he began to teach them. He said:
3 Blessed are the poor in spirit, for theirs is the kingdom of heaven.
4 Blessed are those who mourn, for they will be comforted.
5 Blessed are the meek, for they will inherit the earth.
6 Blessed are those who hunger and thirst for righteousness, for they will be filled.
7 Blessed are the merciful, for they will be shown mercy.
8 Blessed are the pure in heart, for they will see God.
9 Blessed are the peacemakers, for they will be called children of God.
10 Blessed are those who are persecuted because of righteousness, for theirs is the kingdom of heaven.
11 Blessed are you when people insult you, persecute you and falsely say all kinds of evil against you because of me.
12 Rejoice and be glad, because great is your reward in heaven, for in the same way they persecuted the prophets who were before you.
13 You are the salt of the earth. But if the salt loses its saltiness, how can it be made salty again? It is no longer good for anything, except to be thrown out and trampled underfoot.
14 You are the light of the world. A town built on a hill cannot be hidden.
15 Neither do people light a lamp and put it under a bowl. Instead they put it on its stand, and it gives light to everyone in the house.
16 In the same way, let your light shine before others, that they may see your good deeds and glorify your Father in heaven.
17 Do not think that I have come to abolish the Law or the Prophets; I have not come to abolish them but to fulfill them.
18 For truly I tell you, until heaven and earth disappear, not the smallest letter, not the least stroke of a pen, will by any means disappear from the Law until everything is accomplished.
19 Therefore anyone who sets aside one of the least of these commands and teaches others accordingly will be called least in the kingdom of heaven, but whoever practices and teaches these commands will be called great in the kingdom of heaven.
20 For I tell you that unless your righteousness surpasses that of the Pharisees and the teachers of the law, you will certainly not enter the kingdom of heaven.
21 You have heard that it was said to the people long ago, 'You shall not murder, and anyone who murders will be subject to judgment.'
22 But I tell you that anyone who is angry with a brother or sister will be subject to judgment. Again, anyone who says to a brother or sister, 'Raca,' is answerable to the court. And anyone who says, 'You fool!' will be in danger of the fire of hell.
23 Therefore, if you are offering your gift at the altar and there remember that your brother or sister has something against you,
24 leave your gift there in front of the altar. First go and be reconciled to them; then come and offer your gift.
25 Settle matters quickly with your adversary who is taking you to court. Do it while you are still together on the way, or your adversary may hand you over to the judge, and the judge may hand you over to the officer, and you may be thrown into prison.
26 Truly I tell you, you will not get out until you have paid the last penny.

[3] Source: http://www.biblegateway.com/passage/?search=Matthew%205&version=NIV, accessed 25/09/2012.

27 You have heard that it was said, ‘You shall not commit adultery.’
28 But I tell you that anyone who looks at a woman lustfully has already committed adultery with her in his heart.
29 If your right eye causes you to stumble, gouge it out and throw it away. It is better for you to lose one part of your body than for your whole body to be thrown into hell.
30 And if your right hand causes you to stumble, cut it off and throw it away. It is better for you to lose one part of your body than for your whole body to go into hell.
31 It has been said, ‘Anyone who divorces his wife must give her a certificate of divorce.’
32 But I tell you that anyone who divorces his wife, except for sexual immorality, makes her the victim of adultery, and anyone who marries a divorced woman commits adultery.
33 Again, you have heard that it was said to the people long ago, ‘Do not break your oath, but fulfill to the Lord the vows you have made.’
34 But I tell you, do not swear an oath at all: either by heaven, for it is God’s throne;
35 or by the earth, for it is his footstool; or by Jerusalem, for it is the city of the Great King.
36 And do not swear by your head, for you cannot make even one hair white or black.
37 All you need to say is simply ‘Yes’ or ‘No’; anything beyond this comes from the evil one.
38 You have heard that it was said, ‘Eye for eye, and tooth for tooth.’
39 But I tell you, do not resist an evil person. If anyone slaps you on the right cheek, turn to them the other cheek also.
40 And if anyone wants to sue you and take your shirt, hand over your coat as well.
41 If anyone forces you to go one mile, go with them two miles.
42 Give to the one who asks you, and do not turn away from the one who wants to borrow from you.
43 You have heard that it was said, ‘Love your neighbor and hate your enemy.’
44 But I tell you, love your enemies and pray for those who persecute you,
45 that you may be children of your Father in heaven. He causes his sun to rise on the evil and the good, and sends rain on the righteous and the unrighteous.
46 If you love those who love you, what reward will you get? Are not even the tax collectors doing that?
47 And if you greet only your own people, what are you doing more than others? Do not even pagans do that?
48 Be perfect, therefore, as your heavenly Father is perfect.

Suggested reading

You can find information on EModE in all textbooks on the history of English (see Part I of the Bibliography). The period is covered by Volume III of *CHEL*.[4] A fairly concise description of the pronunciation and the grammar of the period – together with some texts – is found in Smith (1999). If you are more deeply interested, Görlach (1991) is still the “classic” work; see also Barber (1997).

[4] CHEL = The Cambridge History of the English language, in six volumes. See the Bibliography for details.

3 Pronunciation: Early Modern English and after

3.1 Introduction

In the previous chapter, we looked at the chief grammatical differences between EModE and PdE. In this chapter, we examine the pronunciation of EModE as well as the most important changes that have taken place since then. By the term "pronunciation" I mean phonological changes (including sound changes or changes in word stress) on the one hand, but I would also like to discuss letter-to-sound rules to some extent, because they might prove useful in explaining some of the orthographic peculiarities of English. The issue of orthography will also be taken up in later chapters, especially when discussing Middle English.

You may wonder why I devote a separate chapter to these questions. The reason is that it was during the EModE period when PdE standard accents, especially British and American, started their independent development, but this is also true for other accents to a large extent (such as Irish or Scottish English, for example). In other words, an understanding of the sound changes that have taken place since EModE – especially those which are restricted to certain accents of English – provides a historical background to the phonological differences that exist between PdE accents. As for orthography, it should be noted that some of the characteristic spelling conventions of PdE became stabilized (or fixed) during the later part of the EModE period (i.e. between the early 17th and the late 18th centuries). Again, these issues will contribute to an understanding of some of the spelling differences between British and American English (such as British *colour, cancelling* vs. American *color, canceling*).

3.2 The sounds of EModE at around 1600

3.2.1 An illustrative sample of EModE pronunciation: the first 12 verses of Matthew 5 from the King James version

In this section, I give a reasonable reconstruction of how EModE around 1600 might have been pronounced. First, I repeat the first twelve verses of Matthew 5 here for your convenience (see (39)), then I present a phonetic transcription (see (40)). It must be emphasized that we have no direct evidence (such as tape recordings) of the pronunciation of EModE, but we *can* tell a lot about it; nonetheless, opinions vary regarding some points.

Also, some words (auxiliaries, conjunctions, etc., such as *shall* and *and*) have both a full (stressed) and a weak (unstressed) pronunciation in PdE: *and* is usually pronounced **ənd**, i.e. weak, unless stressed, in which case it is pronounced with a full vowel (i.e. **ænd**). It seems certain that such a distinction already existed in EModE, but we cannot be sure about the details. Therefore, for the sake of simplicity, I provide the full forms in all cases.

(39) *The first 12 verses of Matthew 5*
1: And seeing the multitudes, he went up into a mountain: and when he was set, his disciples came unto him:
2: And he opened his mouth, and taught them, saying,
3: Blessed are the poor in spirit: for theirs is the kingdom of heaven.
4: Blessed are they that mourn: for they shall be comforted.
5: Blessed are the meek: for they shall inherit the earth.
6: Blessed are they which do hunger and thirst after righteousness: for they shall be filled.
7: Blessed are the merciful: for they shall obtain mercy.

8: Blessed are the pure in heart: for they shall see God.
9: Blessed are the peacemakers: for they shall be called the children of God.
10: Blessed are they which are persecuted for righteousness' sake: for theirs is the kingdom of heaven.
11: Blessed are ye, when men shall revile you, and persecute you, and shall say all manner of evil against you falsely, for my sake.
12: Rejoice, and be exceeding glad: for great is your reward in heaven: for so persecuted they the prophets which were before you.

(40) *A phonetic transcription of the passage in (39)*

1. ænd ˈsiːɪŋ ðə ˈmʊltɪtjuːdz, hiː went ʊp ˈɪntuː ə ˈməʊntɪn, ænd hwen hiː wɒz set, hɪz dɪˈsəɪpəlz kɛːm ˈʊntuː hɪm
2. ænd hiː ˈoːpənd hɪz məʊθ, ænd tɔːt ðem, ˈsɛːɪŋ
3. ˈblesɪd ɑr ðə puːr ɪn ˈspɪrɪt, fɔr ðɛːrz ɪz ðə ˈkɪŋdəm ɒv ˈhevən
4. ˈblesɪd ɑr ðɛː ðæt moːrn, fɔr ðɛː ʃæl biː ˈkʊmfərtɪd
5. ˈblesɪd ɑr ðə miːk, fɔr ðɛː ʃæl ɪnˈherɪt ðɪ ɜrθ
6. ˈblesɪd ɑr ðɛː hwɪtʃ duː ˈhʊŋgər ænd θɜrst ˈæftər ˈrəɪtjəsnəs, fɔr ðɛː ʃæl biː fɪld
7. ˈblesɪd ɑr ðə ˈmɜrsɪfʊl, fɔr ðɛː ʃæl əbˈtɛːn ˈmɜrsɪ
8. ˈblesɪd ɑr ðə pjuːr ɪn hɑrt, fɔr ðɛː ʃæl siː gɒd
9. ˈblesɪd ɑr ðə ˈpeːsˌmɛːkərz, fɔr ðɛː ʃæl biː kɔːld ðə ˈtʃɪldrən ɒv gɒd
10. ˈblesɪd ɑr ðɛː hwɪtʃ ɑr ˌpɜrsɪˈkjuːtɪd fɔr ˈrəɪtjəsnəs sɛːk, fɔr ðɛːrz ɪz ðə ˈkɪŋdəm ɒv ˈhevən
11. ˈblesɪd ɑr jiː, hwen men ʃæl rɪˈvəɪl juː, ænd ˌpɜrsɪˈkjuːt juː, ænd ʃæl sɛː ɔːl ˈmænər ɒv ˈiːvəl əˈgɛːnst juː ˈfɔːlslɪ, fɔr məɪ sɛːk
12. rɪˈdʒɔɪs, ænd biː ɪkˈsiːdɪŋ glæd, fɔr greːt ɪz juːr rɪˈwɔrd ɪn ˈhevən, fɔr soː ˌpɜrsɪˈkjuːtɪd ðɛː ðə ˈprɒfɪts hwɪtʃ wɜr bɪˈfoːr juː

Again, let me emphasize that opinions vary as to some details. Some authors, for example, claim that words such as *mu<u>l</u>titude, <u>u</u>nto, <u>up</u>,* etc., which have an **ʌ** in PdE, and transcribed with an **ʊ** here (= having the same sound as PdE *bush, put,* etc.), were already pronounced with a different vowel, something between **ʊ** and **ʌ**, transcribed with the IPA symbol **ɤ**[5]. This is, indeed, possible – at least in popular London speech; nevertheless, the pronunciation with **ʊ** was certainly possible, especially in formal speech, so I stick to this

[5] In the IPA, this symbol represents a back mid-high unrounded vowel, - like Hungarian *o* but pronounced with spread, i.e. unrounded lips.

variant for the sake of simplicity. Let us now turn our attention to the sound system of EModE at around 1600.

3.2.2 The sounds of EModE (at around 1600) and their typical spellings

Let us start with two important notes. First, the consonants of EModE are the same as those of PdE (with the exception of **ʒ**, as in PdE *vi<u>s</u>ion*, which was coming into existence right then: it probably existed in popular speech already, but not yet in formal speech styles; we will discuss its "birth" later on), so I disregard consonants here, at least for the time being. Second, as far as vowels are concerned, I will not consider unstressed (weak) vowels. The reason for this is that there is much uncertainty about them on the one hand; on the other hand, they are likely to have been mostly identical to PdE ones.This means that I concentrate on stressed vowels. The following table sums up these, then. Since – as I mentioned earlier – you are most likely to read EModE texts with a modernized spelling, I give the examples according to their present-day spelling. Historically, this is sometimes anachronistic (e.g. the letter V was also used where we use U today, as in *vpon*, PdE *upon*); also, EModE texts show considerable variation in spelling (e.g. *cause ~ cawse*). The point is that I wish to illustrate EModE stressed vowels, giving examples, so you can compare them to their modern counterparts and their spellings. It must be emphasized that the typical PdE spellings are also found in EModE, even if not exactly in the same words (so I am not cheating much).

(41) *The stressed (full) vowels of EModE (around 1600)*

EMODE VOWELS	TYPICAL SPELLING(S)	EXAMPLES
ɪ	I, Y	*sit, myth*
e	E	*bless, set*
æ	A	*cat, bad, fast, after*
ʊ	U	*cut, up, full, bush*
ɒ	O	*pot, god, s<u>o</u>rry*
ɔ	O*	*for, lord*
ɑ	A*	*far, part*
ɜ	I*, E*, U*	*sir, first, her, m<u>e</u>rcy, spur, turn*
ɑː	A	*palm, f<u>a</u>ther*
ɔː	AU, AW	*cause, law*
iː	EE, IE, E	*meet, field, he, beer, pierce, here*
eː	EA, E	*peace, meat, fear, s<u>e</u>rious*
ɛː	A, AI, AY, EI, EY	*name, care, rain, day, their, they, fair*
uː	OO	*moon, root, poor, book, stood*
oː	O, OA, OU, OW	*no, load, sore, boar, soul, know, source*
əɪ	I, Y	*time, thy, fire, pyre*
əʊ	OU, OW	*house, town, flour*
ɔɪ	OI, OY	*choice, joy*

Note 1: A * after a vowel letter means that the given letter (O, A, I, E, U) is followed by either (i) a word-final R, cf. *for, far, sir*, etc., or (ii) an R + another consonant (e.g. *lord, part, turn*, etc.).

Note 2: As in PdE, some untypical spellings occur, too. For example, the vowel **ʊ** (as in *cut, up, full*) is sometimes spelt with the letter O, as in *wolf, love, come*.

Note 3: the long vowel **uː** is also found in the sequence **juː**, as in *cute, new, blue, suit, pure*, etc., variably spelt as U, EU, EW, UI. The sequence **juː** derives from an earlier diphthong **iu**, still existing in the early 16th century, but it had probably become **juː** by Shakespeare's time, at least in popular speech. By PdE, the **j** has been dropped in several cases (a process referred to as Yod-Dropping), to be discussed below.

As you can see, the vowels of EModE often coincide with those of PdE, but not always, and there are some variations on the theme. This is the topic of the following sections.

3.3 Vowel changes from EModE to PdE

Let us now repeat the table in (41) as (42), this time including the present British RP pronunciations as well. You will find that sometimes EModE sounds undergo a split (indicated by showing two or more correspondences in the rightmost column), while sometimes they undergo merger, which is shown by identical funny symbols after the modern RP equivalents which originate from distinct EModE ones (for example, the symbol ☎ indicates that the EModE vowels **iː** and **eː** have merged – in a single one, i.e. **ɑː**, in words like *far, part* vs. *palm, father*, which had distinct vowels in EModE).

(42) *EModE sounds, typical spellings, and their typical developments into RP*

EMODE VOWELS	TYPICAL SPELLING(S)	EXAMPLES	RP CORRESPONDENCES
ɪ	I, Y	*sit, myth*	**ɪ**
e	E	*bless, set*	**e**
æ	A	*cat, bad, fast, after*	(a) **æ** as in *cat, bad* (b) **ɑː** as in *fast, after* ☎
ʊ	U	*cut, up, full, bush*	(a) **ʌ** as in *cut, up* (b) **ʊ** as in *full, bush* 💣
ɒ	O	*pot, god, sorry*	**ɒ**
ɔ	O*	*for, lord*	**ɔː** ☠
ɑ	A*	*far, part*	**ɑː** ☎
ɜ	I*, E*, U*	*sir, first, her, mercy, spur, turn*	**ɜː**
ɑː	A	*palm, father*	**ɑː** ☎
ɔː	AU, AW	*cause, law*	**ɔː** ☠
iː	EE, IE, E	*meet, field, he, beer, pierce, here*	(a) **iː** as in *meet, field, he* ✋ (b) **ɪə** as in *beer, pierce, here* ✋
eː	EA, E	*peace, meat, fear, serious*	(a) **iː** as in *peace, meat* ✋ (b) **ɪə** as in *fear, serious* ✋
ɛː	A, AI, AY, EI, EY	*name, care, rain, day, their, they, fair*	(a) **eɪ** as in *name, rain, day, they* (b) **eə** as in *care, their, fair*
uː	OO	*moon, root, poor, book, stood*	(a) **uː** as in *moon, root* (b) **ʊə** as in *poor* (c) **ʊ** as in *book, stood* 💣
oː	O, OA, OU, OW	*no, load, sore, boar, soul, know, source*	(a) **əʊ** as in *no, load, soul, know* (b) **ɔː** as in *sore, boar, source* ☠
əɪ	I, Y	*time, thy, fire, pyre*	(a) **aɪ** as in *time, thy* (b) **aɪə** as in *fire, pyre*
əʊ	OU, OW	*house, town, flour*	(a) **aʊ** as in *house, town* (b) **aʊə** as in *flour*
ɔɪ	OI, OY	*choice, joy*	**ɔɪ**[6]

Let us now look at the individual changes in detail, one by one.

1. **The MEAT-Merger** (indicated by the symbol ✋). This change was already discussed in section 1.4.5. Let me repeat it here for convenience. The EModE vowel **eː**, as in *peace, meat, fear, serious*, became **iː** by about 1700, falling together with original **iː** (as in *meet, field, he, beer, pierce, here*). At this point, pairs such as *meet* and *meat*, pronounced with different vowels in Shakespeare's time, have become homophones. Similarly, *beer* (= **biːr**) now rhymes with *fear* (= **fiːr**). Note that the words *beer, fear, pierce, here*, etc. – where the vowel is followed by **r** – are not yet affected by Breaking at this time (Breaking was also

[6] A triphthong **ɔɪə** is also found in RP, but only in one or two rare words (such as *Moir*), so it is marginal: I disregard it here.

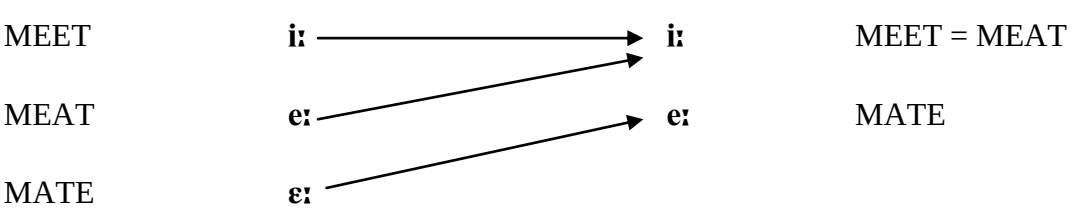
illustrated in Section 1.4.5 and we will give a detailed discussion soon), so *beer* still has a Plain Tense **iː**, just like *meet*.

The MEAT-Merger did not take place in certain varieties of English, most notably, in Irish English, where **eː** remained, so *meet* still sounds different from *meat*, etc. (At least in conservative varieties of Irish English; modern Irish speakers, especially educated people, tend to use **iː** in words like *meat*).

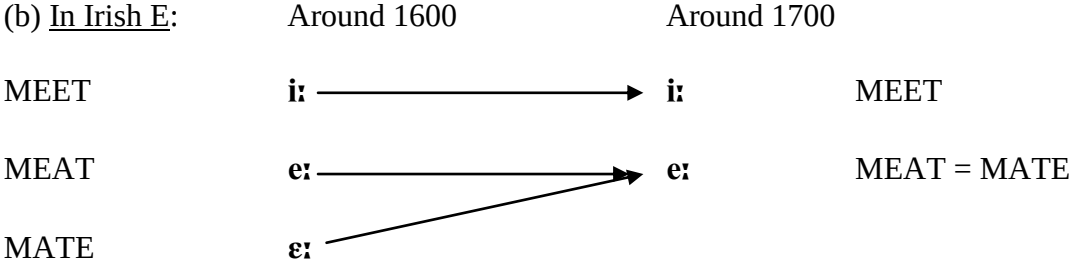
2. EModE **ɛː**, too, underwent a change during the 17th century: it became **eː** by about 1700. So words like *day* and *care*, for example, were pronounced as **deː** and **keːr**, respectively, by the end of the 17th century. This change also took place in Ireland, where, remember, original **eː** remained, so in Irish English, EModE **ɛː** falls together with **eː** - an example is provided by *meat* vs. *mate*, both pronounced **meːt** in Irish English. Schematically:

(43) MEET vs. MEAT vs. MATE

(a) Generally:



(b) In Irish E:



In standard varieties, too, there are some words in which EModE **eː** remains, such as *break, steak*, which would be expected to be pronounced with an **iː** according to the MEAT-Merger, but they have the same vowel as *mate* (so, for example, *steak* sounds the same as *stake*). This is most probably due to **dialect mixture**: some items entered the standard language from (43b)-type dialects. Such cases are not rare at all. Think of some southern dialects of Hungarian, for example, where the vowel spelt *ö* (= IPA **ø**) is found in many cases where standard Hungarian has an *e*, cf. *embör* vs. standard *ember*. Some forms with *ö*, originating from such dialects, have entered the standard, where now you have doublets such as *fel/föl*, *seper/söpör* – in some cases, the form with *ö* has become the normally used one, cf. *sör*, where the form with *e*, i.e. *ser*, sounds old-fashioned or at least rare.

3. The developments of EModE **ʊ**. This vowel generally becomes **ʌ** by the end of the 17th century, as illustrated by *cut, up, sun, much*, but also *love, son, come*, with the same vowel but with an untypical spelling (i.e. with the letter O). In some words, however, **ʊ** remains (as in *full, bush, pull, wolf*). This mostly happens after labial consonants, but not

consistently (cf. *pun*, where – although preceded by a labial consonant, just like in *pull* – **ʊ** does change to **ʌ**).

In fact, this difference may be the result of dialect mixture, too. In the North of England, **ʊ** always remains unchanged, so *much* is pronounced **mʊtʃ**, for example, up to the present day. The Northern dialects may have had some influence on the pronunciation of some words; at any rate, the question remains unsettled.

As indicated by the symbol  , however, unchanged **ʊ** can merge with EModE **uː**, too, so the words *book* (with an EModE long vowel) and *put* (with an EModE short vowel) have the same vowel today. The reason for this is that after the early 17th century, long **uː** (spelt OO) was shortened in some words (such as *book, took, look, stood, good*). This shortening seems to have been regular before **k**, but otherwise it is sporadic, and regularly, no shortening takes place (cf. *stood* with a short vowel but *mood, rood, food* with a long one).

4. **The TRAP-BATH Split**, shown by words such as *cat, bad, fast, after* in (42). The words *trap* and *bath* are also good examples, indeed, these ones are used in the traditional name of the change. The essence of the change is that in EModE, these words all had an **æ**, which, however, changed to **ɑː** by the 19th century in Southern British English in some items, viz. *fast, after, bath*, etc. This change usually happens before voiceless fricatives, but not consistently (cf. *mass* or *math*, still pronounced with an **æ** in RP), and sometimes it also takes place where there is no following voiceless fricative, as in the contracted form *can't*. The irregularities, again, are probably due to dialectal influence: in most of England – indeed, in most of the English-speaking world – the split does not happen. Most notably, it fails to take place in General American, where these words are still pronounced with an **æ**.

5. <u>The diphthongs **əɪ** and **əʊ**</u>. These diphthongs, found in words like *time* and *house*, respectively, around 1600, changed to **aɪ** and **aʊ**, respectively, by about 1700, yielding their PdE standard forms. Words such as *fire* or *flour* were also pronounced with the same vowels: the modern RP pronunciation with a triphthong is the result of Breaking.

6. **Breaking**. As mentioned earlier, this change affects Tense vowels, i.e. (most[7]) long monophthongs and diphthongs, and takes place during the 18th century. Originally, Breaking meant the insertion of a schwa (**ə**) after a Tense vowel when the vowel was followed by **r**; later, however, long monophthongs were **contracted** with the schwa (by the end of the 19th century at the latest), yielding a diphthong. Look at (44) below for the individual developments:

[7] Of the EModE long monophthongs, **ɑː** (as in *calm*) and **ɔː** (as in *law*) were not affected. These two long monophthongs can be collectively described as low (or open), articulatorily speaking, so we can say that only non-low monophthongs underwent breaking (plus, of course, diphthongs).

(44) *Breaking in Southern British English*

Around 1700	Around 1800	Around 1900	Examples
iː	**iːə**	**ɪə**	*beer, pierce, here, fear, serious*
uː	**uːə**	**ʊə**	*poor, cure, sure, fury*
eː	**eːə**	**eə**	*care, their, fair*
oː	**oːə**	**ɔə** 🕯	*sore, boar, source, story*
aɪ	**aɪə**	**aɪə**	*fire, pyre, tire*
aʊ	**aʊə**	**aʊə**	*flour, sour, hour*

As you can see, the diphthongs remain unchanged (see the shaded rows), but the long monophthongs show contraction. As mentioned earlier, the diphthong **ɔə** has become monophthongized to **ɔː** in RP (and in Southern British English in general); this diphthong is shown by the symbol 🕯 in the table. Otherwise, the situation we find at about 1900 is the same as in present-day RP (at least conservative RP, because other Broken Tense vowels are also affected – to varying degrees – in more innovative speech, especially the speech of young generations; this, however, is a topic that is of concern to PdE phonology and dialectology).

It must be noted that Breaking varies a lot across English accents. It is typically found in Non-Rhotic accents, i.e., ones which also exhibit R-Dropping, but there is no direct correspondence. In Scottish English, for example, Breaking is practically unknown. In American English, it is variable: sometimes you have Breaking, but the effect of a following **r** may also manifest itself as the changing of the Tense vowel into a Lax one, so *beer* may be pronounced as **bɪər** or **bɪr**, etc., but – since there is no R-Dropping in general American English, Broken Tense vowels are still allophones of Plain Tense ones, as they indeed were in England, too, before R-Dropping took place (cf. Section 1.4.5). In Welsh English (which is Non-Rhotic, just like RP), you do have Breaking, but Broken Tense vowels do not undergo contraction, so *beer* is pronounced **biːə**, and *pure* is pronounced **pjuːə** in Welsh English, for example.

7. "Mid-high diphthongization". The mid-high (or, to use another phonetic term, half-close) long monophthongs, i.e. **eː** and **oː**, become **eɪ** and **oʊ**, respectively, during the 19th century, unless undergoing Breaking before. That is, this is their normal development when not followed by **r**. Look at table (45) below:

(45) *Mid-high diphthongization*

18th century	19th century	PdE RP	Examples
eː	**eɪ**	**eɪ**	*name, rain, day, they*
oː	**oʊ**	**əʊ**	*no, load, soul, know*

This process is not found in Scottish, Irish, or Welsh English, as well as in parts of England, and several North American speakers also pronounce a monophthong. As you can see in (45), the diphthong **oʊ** underwent a further shift to **əʊ** in Southern British English, including RP.

8. You will have noticed that some of the changes shown in (42) above have not been discussed yet. These are the ones indicated by the symbol * in (42), repeated here (46) for your convenience:

(46)

EMODE VOWELS	TYPICAL SPELLING(S)	EXAMPLES	RP CORRESPONDENCES
ɔ	O*	*for, lord*	**ɔː**
ɑ	A*	*far, part*	**ɑː**
ɜ	I*, E*, U*	*sir, first, her, mercy, spur, turn*	**ɜː**

Remember that * after a vowel letter means that the given letter (O, A, I, E, U) is followed by either (i) a word-final R, cf. *for, far, sir*, etc., or (ii) an R + another consonant (e.g. *lord, part, turn,* etc.). This also means that the EModE vowel sounds found in this table are always followed by a pronounced **r**, while in PdE RP – and other Non-Rhotic accents – they are not (except if the **r** is word-final, where it does appear as a Linking-R if the next word begins with a vowel). The reason for this is R-Dropping, a change in the distribution of **r**, to which we now turn.

3.4 Consonant changes from EModE to PdE

3.4.1 R-Dropping (Non-Rhoticity) and related vowel changes

3.4.1.1 R-Dropping itself

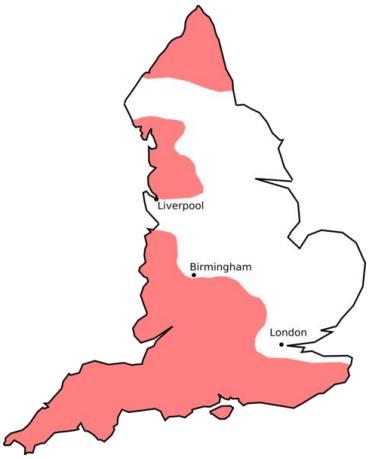
During the 18th century, **r** was becoming gradually weakened and finally dropped before consonants as well as word-finally (cf. the words given in (46) above, or the passage from the King James Bible, transcribed, in (39) and (40)), pronounced with a **r** in EmodE but without it in many accents of English – such as RP – today). Accents of English which have R-Dropping are called **non-rhotic**, while those which lack it are called **rhotic** (such as most American accents). The phenomenon of R-Dropping is also referred to as **non-rhoticity**, therefore. R-Dropping started as a feature of popular speech in and around London, but it soon became quite widespread, in two ways:

1. It spread geographically, to most of parts of England, as well as to Wales, some areas of the US (but not to Canada), and the Englishes of the southern hemisphere (= South Africa, New Zealand and Australia). Its spread in England was of course gradual, as illustrated by the following maps below[8]:

[8] Source: http://en.wikipedia.org/wiki/Rhotic_and_non-rhotic_accents, accessed 29/09/2012.

(47)

(a)



(b)

On the first map, in (47a), the rhotic areas of England are shown (shaded areas) at around 1950, while the shaded parts in (47b) show the rhotic areas in the late 20th century. As you can see, rhoticity is on the retreat in England: this is because by the middle of the 19th century, R-Dropping (non-rhoticity) had become the standard, educated form of pronunciation in the land, adopted by more and more speakers. I do not attempt to give a detailed discussion of rhotic vs. non-rhotic accents here: it is the subject of dialectology, really. Nonetheless, I present another map, showing the traditional non-rhotic areas of the USA, most of which, remember, is rhotic:

(48)[9]

The map in (48) shows the early 21st century situation, among white speakers (African American speech is mostly non-rhotic in other parts, too).

[9] Source: http://en.wikipedia.org/wiki/Rhotic_and_non-rhotic_accents, accessed 29/09/2012.

2. While originating in popular speech, non-rhoticity spread socially, too: in England (as mentioned), it became more and more accepted in educated circles and higher levels of society, so by the middle of the 19th century, it had come to be regarded as the educated, generally accepted form of pronunciation.

Rhoticity has generally remained in the Celtic countries, i.e. Scotland and Ireland, with the exception of Wales. This is because, in the early 19th century, most people in Wales still spoke Welsh rather than English. The majority of Welsh speakers learnt English at school after compulsory primary school education was introduced in the late 19th century, by which time R-Dropping had become the norm in England, so it was this (non-rhotic) way of pronunciation which was taught to Welsh children by English teachers. In Scotland or Ireland, however, English had been spoken much more extensively for centuries. In the US, the situation is complex, but I would like to give one example: in New England (= the North-East), which had close commercial and cultural connections with Britain, the emerging new pronunciation was considered to have a high social prestige – a fact which undoubtedly contributed to these areas becoming non-rhotic. Finally, as far as the southern hemisphere is concerned, note that these areas underwent massive colonization by Britain during the later part of the 19th century, by which time non-rhoticity had become the norm in England. Also, most English emigrants to South Africa, New Zealand and Australia arrived from the South-East of England during the first phase of emigration, so they spoke a non-rhotic accent.

It is important to note again that R-Dropping only takes place if the **r** is (i) word-final, or (ii) it is followed by another consonant. So, you have no R-Dropping in words like *red, true, fairy, merry*, where the **r** is followed by a vowel sound. Note that silent vowel letters – as in *car<u>e</u>*, for example – do not count: in pronunciation, they are not there, so R-Dropping does take place.

Word-final **r**, though generally dropped, behaves specially: it is still pronounced if the next word begins with a vowel. So, *care* is pronounced without a **r** in *I don't care* or *I care for you* (where it is either utterance-final or the next word begins with a consonant): specifically, it is pronounced **keə**. However, its pronounciation is **keər** in *I don't care about it*, because the following word (= *about*) begins with a vowel. This phenomenon is called **Linking-R**, found in most non-rhotic accents of English (but not in all of them: the non-rhotic accents of the Southern USA, for example, have no Linking-R, so word-final **r** is always dropped). Linking-R has given rise to another phenomenon, called **Intrusive-R**, to be discussed below. Before that, however, an important consequence of R-Dropping should be described, concerning short vowels before a dropped **r**, as illustrated by the examples in (46).

3.4.1.2 The lengthening of short vowels before a dropped *r*

Let me repeat the table in (46) here for your convenience as (49):

(49)

EMODE VOWELS	TYPICAL SPELLING(S)	EXAMPLES	RP CORRESPONDENCES
ɔ	O*	*for, lord*	**ɔː**
ɑ	A*	*far, part*	**ɑː**
ɜ	I*, E*, U*	*sir, first, her, m<u>e</u>rcy, spur, turn*	**ɜː**

As you can see, these EModE vowels, when followed by **r**, were pronounced short. In RP, however – and indeed, in non-rhotic accents in general – they have a long vowel. So, for example, *lord* was pronounced **lɔrd** in EModE, as it still is in rhotic accents of English, but it is pronounced **lɔːd** in RP and in other non-rhotic accents. The same goes for words such as *far* with a word-final **r** in EModE (= **fɑr** vs. **fɑː**). In other words, if the dropped **r** is preceded by a short vowel, the vowel becomes long[10]. The question is why this happens.

The answer is provided by the syllabic position of the dropped **r**: if a **r** is word-final, or if it is found before a consonant, it is syllable-final, in other words, it is in a **coda** (the term **coda** refers to consonants at the end of a syllable). We can now reformulate R-Dropping as follows:

(50) *R-Dropping revisited*
In non-rhotic English accents, **r** is dropped if it is in a coda.

If the **r** is followed by a vowel, it is not in a coda: instead, it sits in the **onset** of a syllable, a term referring to the consonant(s) found in syllable-initial position. For example, the **r** is in an onset in *fairy*, syllabified as *fai.ry* (where the dot indicates a syllable boundary): it is initial (= an onset) in the second syllable of the word, so it is not dropped. On the other hand, **r** is in a coda position in *fair* or *scarce*, so it is dropped.

The loss of **r** in codas would, however, result in a loss of the overall "length" of the syllable: for example, *card*, pronounced **kɑrd**, would become shorter (= ***kɑd**). The vowel is lengthened, as it were, to compensate for the loss of the **r**: since a long vowel equals a sequence of a short vowel + a consonant in terms of length (or quantity), the form **kɑːd** counts as long as **kɑrd**. This phenomenon is called **compensatory lengthening** (I hope the term is now easy to understand), and it is quite frequent in the world's languages. In many non-standard varieties of Hungarian, for example, **l** is dropped before a consonant (= in a coda, though not always when word-final), so that in such varieties, standard *volt* or *zöld* appear as *vót* and *ződ*, respectively: note that the vowel is lengthened!

You may now ask why there is not compensatory lengthening in words where the vowel had been long before R-Dropping, i.e. where there is a Broken Tense vowel, as in *beard*, pronounced **bɪərd** before R-Dropping, but **bɪəd** afterwards. The reason is that the diphthong **ɪə** already counts as a long vowel, and vowels are either short or long: in other words, long vowels cannot be lengthened to "overlong" vowels. This fact blocks compensatory lengthening. The same situation is observed in varieties of Hungarian mentioned above: if the vowel is originally long before the dropped **l**, no further lengthening takes place, cf. *csinált*, pronounced *csinát* in L-Dropping accents, rather than with an overlong **csináát*.

A further important consequence of compensatory lengthening is that the lengthened **ɔ** (as in *for, lord*) falls together, i.e. merges, with the EModE vowel **ɔː**, as in *law*, *cause*, etc. Since the Breaking of EModE **oː** (as in *sore* or *source*, cf. (44) above) also produces **ɔː** in PdE RP, the modern RP vowel **ɔː** derives from three historically different vowels which have all merged into one vowel in this accent (but not in all accents of English). This explains the problem regarding the Janus-faced behaviour of **ɔː** in RP – see, for example, Nádasdy (1996) for details.

[10] Strictly speaking, this is only true for stressed vowels. In *better*, for example, pronounced ˈ**betər** in EModE (and in rhotic accents), the schwa isn't lengthened in non-rhotic accents (cf. RP ˈ**betə**), because it is unstressed. See also *shepherd*, pronounced ˈ**ʃepəd** in RP, for the same reason. In fact, schwa cannot be long in general.

3.4.1.3 The appearance of Intrusive-R

In Later ModE, becoming widespread in the 20th century, another change took place in most non-rhotic accents (including RP), referred to as **Intrusive-R**. It means that in certain cases, an **r** is inserted at the end of a word where historically, there had been none, if the word is followed by another word which begins with a vowel. Examples include *law[r] and order*, *the spa[r] is nice*, *vanilla[r] ice*, etc., where the bracketed [r]'s indicate an **r** which is pronounced, though it wasn't originally present (and neither is it represented in writing). Intrusive-R appeared on the analogy of Linking-R. Let us see the details.

For historical reasons (some of which have been discussed already), **r** in word-final position was found only after certain vowels, but not others, at the time when R-Dropping took place (= around 1800):

1. Breaking is one reason: before **r** (including a word-final **r**), Plain Tense vowels were replaced by Broken Tense ones, so Plain Tense vowels were no longer found in that position, cf. *bee* vs. *beer*, pronounced **biː** and **bɪər**, respectively, before R-Dropping. After R-Dropping, word-final **r** was still retained as a Linking-R if the following word began with a vowel, as in *the bee**r** is good*. Remember that originally, all Broken Tense vowels ended in a schwa, and, indeed,

2. a final unstressed sequence **ər**, as in *better*, also became **ə** upon R-Dropping, but again, the **r** was retained as a Linking-R, as in *bette**r** idea*.

3. Finally, short full (stressed) vowels have never been possible in word-final position in English. Indeed, this is a characteristic property of all Germanic languages: word-final vowels are either long (and then stressed) or unstressed (but then, they are short). To put it simply, this means that no short vowel, with the exception of schwa (which is always unstressed) is possible in word-final position, so no English word ends in **e, ɑ, ɜ, ɔ, ɒ, æ, ʊ, ʌ**, which are all full vowels[11]. Furthermore, due to reasons that need not concern us here at this point, the only short stressed vowels found before a word-final **r** are **ɑ, ɔ, ɜ** in EModE, cf. (49) above.

All in all, at the time of R-Dropping, only a limited group of vowels was possible before a word-final **r**, as shown by the examples in (51):

[11] For the distinction between full and weak vowels, see, for example, Nádasdy (1996). The vowel **ɪ** is a notable exception, as in *city*, i.e. **ˈsɪtɪ**, which does occur word-finally, but in such cases, it is always unstressed (weak); indeed, it is often replaced by long **iː** in this position. At any rate, unstressed **ɪ** was never followed by word-final **r** at the time of R-Dropping, anyway.

(51) *Possible word-final "vowel + **r**" sequences at the time of R-Dropping*

Vowel Types	Vowels + r	Examples
Short stressed vowel	**ɑr**	*far, star*
	ɔr	*nor, for*
	ɜr	*sir, her, spur*
Broken Tense vowel (= long stressed vowel before **r**)	**ɪər**	*beer, here*
	ʊər	*poor, sure*
	eər	*care, fair*
	ɔər	*sore, boar*
	aɪər	*fire, pyre*
	aʊər	*hour, flower*
Unstressed vowel (= schwa)	**ər**	*bett<u>er</u>, bak<u>er</u>*

If you study the table in (51) carefully, you will notice that word-final **r** is found *exclusively after non-high vowels*, but *never after high ones*, including (i) the long monophthongs **iː** and **uː**, which are high monophthongs, (ii) any diphthong whose second element is **ɪ** or **ʊ** – these second elements are high vowels themselves, cf. **aɪ, eɪ, oʊ, aʊ**.

Nevertheless, R-Dropping results in some changes which alter the picture to some extent. The table above is repeated here as (52), but this time, showing the situation *after* R-Dropping as well as the compensatory lengthening of short stressed vowels (which went hand in hand with R-Dropping):

(52) *Possible word-final vowels after R-Dropping*

Vowel Types	Vowels + r	Examples
Long stressed monophthong	**ɑː**	*far, star*
	ɔː	*nor, for*
	ɜː	*sir, her, spur*
Broken Tense vowel (= a stressed diphthong ending in a schwa)	**ɪə**	*beer, here*
	ʊə	*poor, sure*
	eə	*care, fair*
	ɔə 💣	*sore, boar*
	aɪə	*fire, pyre*
	aʊə	*hour, flower*
Unstressed vowel (= schwa)	**ə**	*bett<u>er</u>, bak<u>er</u>*

Note that the vowels **ɑː** and **ɔː** (see the shaded boxes in (52)) had already existed in EModE – crucially, when *not* followed by **r** – check the table in (52), itself a reproduction of a part of the table in (41) above:

(52)

Vowel	Typical spelling	Examples
ɑː	A	*palm, father*
ɔː	AU, AW	*cause, law*

Although **ɑː** was rare in word-final position in EModE, it did exist, as in the words *shah* or *spa*[12], and later on, it became more and more frequent due to the appearance of new words such as *bra, blah*, etc. The vowel **ɔː** was more frequent word-finally, witness *law*, *saw* 'fűrész' and *saw* 'the past tense of *see*', *raw*, *paw*, *draw*, to mention just some examples. Furthermore, the monophthongization of Broken Tense **ɔə** to **ɔː** (as in *sore, bore*, etc., as shown by the symbol 💣 in table (51), cf. also (44) and the discussion after it), becoming widespread in Southern British English in the early 20th century, also increased the occurrence of **ɔː** in word-final position; in such words, however, there *was* a word-final **r** originally (the diphthong **ɔə** results from Breaking, remember).

Similarly, final schwa also existed at the time of R-Dropping, as in the words *errata, extra, vanilla, idea*, etc. (note that these words are typically Latin or Romance loans). In most cases, however, schwa was typically followed by a **r** at the end of words, as in *better, water, hammer, master, colour, linear, Tudor*, etc.; note the variety of spellings for word-final **ər**. To sum up, non-high vowels in word-final position were mostly followed by **r**, and the number of words in which they were not definitely formed a minority.

The point is that, after R-Dropping, as well as the monophthongization of **ɔə** to **ɔː**, many words which had earlier been distinguished by having vs. not having a **r** at the end – such as *saw* vs. *sore* (both = **sɔː** in popular speech by the early 20th century when pronounced in isolation[13]) became homophones. The same is true for pairs such as *diner – Dinah* (both pronounced **ˈdaɪnə** in isolation), or *spa – spar* (= **spɑː** in isolation). Furthermore, stressed **iː** before a schwa (as in *idea*, pronounced **aɪˈdiːə**) has generally undergone contraction, producing the new form **aɪˈdɪə**, just like vowel + schwa sequences resulting from Breaking (cf. the table in (44) in Section 3.3. above), so *idea* may now rhyme with *dear*. (It must be added that contraction is still optional in words like *idea*, so **aɪˈdiːə** and **aɪˈdɪə** are both possible, though the contracted variants are more common nowadays; as opposed to this, contraction has always taken place after Breaking, so *dear* is never pronounced ***diːə** in RP as well as in most non-rhotic accents[14]).

Now, as I mentioned, words with an original **r** at the end could still retain the **r** as a Linking-R, so R-ful and R-less variants have come into existence: *dear* is pronounced **dɪə** in isolation, or when sentence-final, or if the next word began with a consonant (e.g. *Oh, my dear!* or *my dear friend*), but as **dɪər** in *dear Ann*, for instance. Similarly, *sore* is pronounced **sɔː** (as in *sore throat*) but **sɔːr** in, e.g., *sore is*. On the analogy of Linking-R, then, Intrusive-R was born: words ending in a non-high vowel but originally having no **r** at the end, started to be pronounced with a final **r**, too, if the next word began with a vowel. Examples include (where the capital R indicates an Intrusive-R) *I sawR it, lawR and order, the shahR of Persia,*

[12] A loanword originating from the name of a Belgian resort which had long been popular. May I note that this long vowel was relatively rare in general in EModE before the TRAP-BATH Split.

[13] The term *in isolation* means that the word is pronounced alone, not within a sentence (like when you quote words).

[14] Remember that Welsh English is an exception, where contraction hasn't taken place.

the spaR is nice, the ideaR is clear, vanillaR ice, etc. Intrusive-R means therefore a **r** which was not historically present: this is also shown by the fact that it is not shown in spelling. An interesting consequence of it is that words which used to differ in the presence vs. absence of a final **r** only, such as *saw – sore, law – lore, spa – spar*, etc., are now *always* homophones: they are all pronounced with a final **r** if the next word begins with a vowel but without a final **r** otherwise.

Intrusive-R is, of course, not found in rhotic accents, where R-Dropping has not taken place, or in those non-rhotic accents (such as the ones in the South-East of the US) where there is no Linking-R. This is understandable: in both types of accent, Linking-R is missing, so there was no source of analogy for creating an Intrusive-R.

As I said, Intrusive-R is a relatively recent phenomenon, and even a few decades ago, it was regarded upon quite negatively by many speakers, but nowadays it is very common, and most speakers – even in RP – tend to use it.

3.4.2 Palatalization

In EModE, at around 1600, short unstressed **ɪ** changed – optionally – to **j** (a glide called Yod in phonetics) before a following unstressed vowel, i.e. a schwa. This variation still exists in many cases, as in the word *opinion*, which can be pronounced both as **əˈpɪnɪ̲ən** as well as **əˈpɪnjən**. In Shakespeare's time, this was also the case in words like *vision*, variably pronounced as **ˈvɪzɪ̲ən** or **ˈvɪzjən**, or in *special*, pronounced as **ˈspesɪ̲əl** or **ˈspesjəl**. Note, however, that these words are pronounced different today, i.e. as **ˈvɪʒən** and **ˈspeʃəl**, respectively. The reason for this is Palatalization, a process affecting alveolar obstruents, that is, **s, z, t, d**, before a Yod, resulting in palatal obstruents, according to the following pattern:

(53) **s** > **ʃ**
z > **ʒ**
t > **tʃ**
d > **dʒ**

The Yod itself disappears in the process. Look at the following table for examples:

(54)

Before Palatalization	After Palatalization	MoE (RP)	MoE spelling
ˈvɪzjən	**ˈvɪʒən**	**ˈvɪʒən**	*vision*
ˈmezjər	**ˈmeʒər**	**ˈmeʒə**	*measure*
ˈspesjəl	**ˈspeʃəl**	**ˈspeʃəl**	*special*
ˈnɛːsjən	**ˈnɛːʃən**	**ˈneɪʃən**	*nation*[15]
ˈnɛːtjər	**ˈnɛːtʃər**	**ˈneɪtʃə**	*nature*
ˈkwestjən	**ˈkwestʃən**	**ˈkwestʃən**	*question*
ˈsoːldjər	**ˈsoːldʒər**	**ˈsəʊldʒə**	*soldier*

These words, of course, were also pronounceable with an **ɪ** instead of the Yod, cf. the variant Shakespearean pronunciations of *vision* and *special* above. In such cases, no Palatalization takes place (since there is no Yod to cause it), but interestingly, these variants gradually disappear, so forms like ˈ**vɪzɪən** (for *vision*), etc., are no longer found, although in more recent or rare, especially formal words, such as *radiant*, for example, Palatalization has never taken place, at least in standard English. Other consonants do not undergo Palatalization (cf. *opinion, salient*).

Yod, however, is also found in the sequence **juː**, when it is unstressed (the **uː** may itself become short in certain cases, as in *annual* (= ˈ**ænjuːəl** or ˈ**ænjʊəl**)[16]. If the **juː** is preceded by **s, z, t, d**, it often causes Palatalization, too. The word *gradual*, for example, can be pronounced as ˈ**grædjuəl** or ˈ**grædʒuəl**, although the palatalized variants are becoming more and more frequent; the word *sexual*, for example, originally ˈ**seksjuəl**, is hardly ever pronounced like that nowadays, the palatalized variant (=ˈ**sekʃuəl**) having been generalized.

A final note on Palatalization: as I mentioned, it typically occurs in unstressed syllables only, but not in stressed ones, cf. *dune, tune, suit*, etc., pronounced with a stressed **juː** in EModE. So *suit* did not become ***ʃuːt**, for instance. (But the Yod could be dropped later on; see below.) A notable exception is the word *sure*, pronounced **sjuːr** in Shakespeare's time, but with an **ʃ** (e.g. RP **ʃʊə**) in PdE. The interesting point here is that the letter S is practically never pronounced like that in word-initial position, the sound **ʃ** being usually spelt SH in such cases, cf. *shoot, shore, shine*, etc. This irregular letter-to-sound correspondence in *sure* is explained by the exceptional Palatalization found in this word. (Another example is the word *sugar*; I do not know of any other examples.)

[15] The spelling of *nation* is misleading: the *t* would suggest an original **t**, rather than **s**. This word, however, *did* have a **s**, and the spelling with *t* is used as an imitation of the Latin form (= *natio*). Indeed, in Middle English, it was spelt *nacioun*. There are many other examples, e.g. *ration, fiction, deletion, intention, prohibition*. The original **s** is always recoverable from the pronounced form having a **ʃ**, rather than **tʃ**, which is the result of the Palatalization of **t**. Compare *question*, where the *t* represents original **t**.

[16] Some dictionaries, such as the *Longman Pronunciation Dictionary*, use the symbol **u** in such cases. I will follow this practice here. As for when this shortening takes place, see, for example, Nádasdy (1996).

3.4.3 Yod-Dropping

As mentioned in Section 3.2.2. above, there is a sequence **juː** in EModE, as in *cute, new, blue, suit, pure*, etc., variably spelt as U, EU, EW, UI. The sequence **juː** derives from an earlier diphthong **iu**, still existing in the early 16th century, but it had probably become **juː** by Shakespeare's time, at least in popular speech. By PdE, the **j** has been dropped in several cases, a process known as Yod-Dropping. It started in the 17th century and has been gradually spreading ever since; in some cases, it takes place optionally, while in other cases, it occurs in some accents of English but not in others. There are, therefore, two basic phases in the history of Yod-Dropping:

1. **Early Yod-Dropping**, taking place before the 18th century, and affecting all accents of English, and (as a result) the affected words are now invariably pronounced without a Yod, i.e. with simple **uː** rather than **juː**. Look at the following table in (55) for an illustration:

(55) **Early Yod-Dropping**

	Around 1600	**18th century (and today)**	**MoE spelling**
(a)	**tʃjuː**	**tʃuː**	*chew*
	dʒjuːn	**dʒuːn**	*June*
(b)	**rjuːd**	**ruːd**	*rude*
(c)	**bljuː**	**bluː**	*blue/blew*

In (a), the Yod stands after a palatal consonant; in (b), there is a **r** before it, while in (c), it is preceded by a cluster of a consonant + **l**. Early Yod-Dropping always takes place in such cases, cf. also *Jew, rural, glue*, etc., pronounced without a Yod.

2. **Later Yod-Dropping**, starting probably not earlier than 1700, and going on ever since, but affecting various accents of English in different ways. Let us now see when it occurs.

(a) After a single **l**, as well as after **s, z**, it is optional in RP, e.g. in *Luke, suit, Zeus*, pronounceable both with or without a Yod (e.g. **sjuːt** ~ **suːt**), although the yodless pronunciation is now definitely more frequent, especially after **l** (so **ljuːk** for *Luke* sounds old-fashioned). American English, however, always has Yod-Dropping in these cases.

(b) Furthermore, American English also has Yod-Dropping after **t, d, n, θ**, as in *tune, due, new, enthusiasm*, which are pronounced with a Yod in RP.

It must be noted that in all these cases, Later Yod-Dropping is restricted to stressed syllables: even in American English, words like *venue, value*, where the **juː** is unstressed, are pronounced with a Yod!

The process of Yod-Dropping, then, explains why the same typical spellings, i.e. U, EU, EW, UI, are pronounced with a Yod in some cases but without it in others. The variability of Later Yod-Dropping explains both variants found in RP as well as the relevant differences between British and American pronunciation.

3.4.4 WH-Simplification

In most accents of PdE, words spelt with an initial WH, such as *which, when, what*, etc., are generally pronounced with a simple **w**, just like words written with W, so *which* sounds the same as *witch*. In EModE, however, these WH-words were pronounced with initial **hw**. This cluster, however, was simplified to **w** later on in most English accents – although not everywhere: in Scottish English, for example, it has remained up to the present, and many North American speakers also use it.

3.5 Word stress

This discussion is aiming at illustration, rather than a complete discussion, since word stress is a complex issue; also, it has continuously been changing, since as early as the Middle English period, and at all times ever since, it showed a great amount of variation.

As an illustration, I would like to use one example only. Word stress in EModE often fell on the last syllable of words which are stressed on their first syllable today. The most characteristic and frequent example is provided by verbs of three syllables with a long vowel or diphthong in their last syllable, such as *decorate, persecute, signify*, etc. They were, as mentioned, stressed on their last syllable in EModE, i.e. as *decoráte, persecúte, signifý*, etc. Gradually, however, their stress moved to the first syllable, so today, they are stressed as *décorate, pérsecute, sígnify*. Even as late as the 19th century, however, the old stress pattern still existed in variation with the new one, and in Irish English, it still survives. It is interesting that in nouns derived from some of these verbs, such as in *decorátion*, the old place of stress has remained up to the present.

Sometimes nouns (and adjectives) also undergo the same process. Let me use the first verse of William Blake's well-known poem, *The tiger*, from 1794:

(56)

TIGER, tiger, burning bright
In the forests of the night,
What immortal hand or eye
Could frame thy fearful symmetry?

Blake rhymes *eye* with *symmetry*, which is very strange to the modern reader, since the word is pronounced **ˈsɪmətrɪ** today, whereas *eye* is pronounced **aɪ**. In Blake's time, however, it could also be stressed on the last syllable, pronounced **sɪməˈtraɪ**! When the stressed moved to the first syllable, and the final vowel became unstressed, it got reduced to weak **ɪ**, hence the modern form, which no longer rhymes with *eye*. I must add that the modern form must have existed at the end of the 18th century, and it is possible that the old form used by Blake was already archaic (but it was still possible): the situation is similar to the use of *thy*, which, remember, was also archaic by that time! I hope that this shows – like most of this chapter as well as the previous one – is that studying the history of a language is not only an interesting subject in itself, but it also helps in understanding and appreciating earlier literature.

Sometimes words of two syllables also undergo stress shift, and this process may also show dialectal variation. The word *address*, for example, used to be stressed as *addréss*; it still is in British English, but in American, it is *áddress* when used as a noun (still *addréss*

when used as a verb). Conversely, the verb *haráss* is still stressed like that in America, but has undergone stress shift in British (= *hárass*).

This concludes our discussion of the phonological changes between EModE and our day. Let us now make a big jump, going back to the earliest known periods in the history of English.

Suggested reading

You can find information on EModE pronunciation and later developments in most textbooks on the history of English (see Part I of the Bibliography). If you are more deeply interested in EModE pronunciation, see Görlach (1991), Barber (1997) or Volume III of *CHEL*; though not so recent – and rather technical -, Dobson (1957) is still a classic. For sound changes after the late 18th century, see Volumes IV and V of *CHEL*; the latter discusses the development of English everywhere except North America in the past two centuries; Volume VI covers the development of American English.

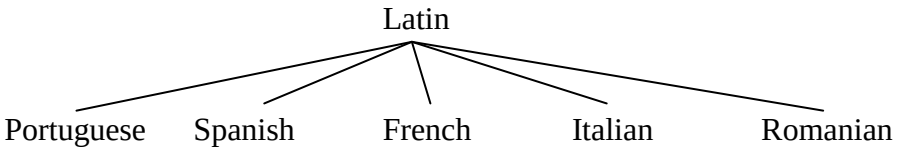
4 The Indo-European family and Proto-Indo-European

4.1 Introduction

In the previous two chapters, we have seen how different changes can lead to differences between varieties (dialects) of a language. The differences between present-day British and American English, both in grammar and pronunciation (and, of course, in vocabulary), are the results of the fact that several changes have taken place in Britain, but not in America, or the other way round. Sometimes, as in the case of the use of tenses, EModE showed variants, with one variant becoming general in America but the other one in Britain. We might as well say that EModE is the ancestor or parent, of which standard British and American are descendants. Similarly, the differences between other varieties – Scottish, Irish, etc. – are mostly due to their different developments since EModE, but it must be added that much of the variation found in the British Isles goes back to earlier times – some as early as the Old English period.

In the case of English, the differing changes have not been significant enough to result in mutual unintelligibility, but it can happen that – given enough time – deviating developments can become so significant that completely different languages come into existence. A nice illustration is provided by Latin, which, after the collapse of the West Roman Empire, developed in various ways in different territories, so it became split into several independent languages (called Romance or Neo-Latin languages), such as Portuguese, French, Spanish, Italian, Romanian, etc. A parent language is often called a **proto-language** in historical linguistics. The descendants of a parent language are called its **daughters**, which are **sisters** to each other and they constitute a family; also, members of the same family are said to be **related**. Families are traditionally represented in the form of a **family tree**, as shown by the (simplified) tree of the Romance family below:

(57)



English is a member of the Germanic family, related, for example, to German, Dutch, Swedish, Norwegian, etc., all descendants of an ancient language called **Proto-Germanic**. We are not going to present a family tree of the Germanic family here, because it will be studied in detail in the next chapter.

However, Germanic itself is a member of an even larger family, called **Indo-European** - to which the Romance languages also belong, as well as many other ones, e.g. Slavonic languages such as Czech, Polish, Bulgarian, etc., or Celtic languages such as Welsh or Sottish Gaelic, etc. The Indo-European languages are all descendants of a language referred to as **Proto-Indo-European**, spoken about 6000 years ago, presumably in Eastern Europe – roughly, north of the Black Sea, according to the most common view. Then, its speakers migrated in different directions, getting separated from each other; as a consequence, the language changed in different ways, slowly giving rise to separate languages, e.g. Proto-Germanic, Proto-Celtic, etc., the daughters of Proto-Indo-European. Then, in turn, these languages split up into several ones in most cases (but not always: as opposed to Germanic, which includes several languages, Hellenic, i.e Greek, for example, has remained as one language). Let us now see the branches of the Indo-European family tree in detail.

4.2 The branches of Indo-European

Indo-European languages as indigenous ones are spoken in most of Europe as well as in South-Western Asia, extending to the Indian subcontinent (hence the term "Indo-European"). As a result of the colonization of other parts of the world, however, they are now spoken on every continent. I will concentrate on the areas where Indo-European languages had been spoken before European colonization, but in the case of the most widespread languages of European colonizers, I also mention the most important places to which the given language was carried via colonization. I start the discussion with the Indo-European languages of Europe, then going on to the Asian ones (so, basically, I proceed from West to East). Also, I concentrate on the present-day languages, generally omitting the earlier stages (or just mentioning them without any detailed presentation). Students who are interested in these aspects are referred to the works given in the Suggested Reading section at the end of this chapter.

4.2.1 Germanic

Germanic is where English belongs, and it will be described in detail in the next chapter, so I am not discussing it here.

4.2.2 Celtic

Celtic (pronounce ˈ**keltɪk**[17]) was spoken over large territories of Europe in Antiquity, extending from the British Isles to Southern and Central Europe, but today, Celtic languages are restricted to North-Western Europe, and they are spoken by a relatively small number of people. They are of some interest, however, to students of English, because most of them are spoken in Britain and Ireland, so they will be discussed in some detail. Modern Celtic languages are divided into two groups.

1. **Brythonic** is represented by two living languages. The first one is **Welsh**, spoken mostly in Wales, but scattered speakers are, of course, found all over Britain (and North America); more interestingly, there is a Welsh community of a few thousand people in Patagonia, in the South of Argentina, which was a popular target area for Welsh emigrants for some time. In Wales, it is spoken by about 600,000 people according to the latest census, i.e. by about 20% of Wales' population. This does not seem much, since the overwhelming majority of Wales speaks English as a native language, but appearances are deceptive. Specifically, Welsh is still a mojority language in the West and the North-West of Wales, especially in the North-West, where it is spoken by up to 90% of the population! The chief reason why English is spoken by about 80% of the entire population of Wales is that the South of Wales is overwhelmingly English-speaking, and this is the most densely populated area. Welsh is now one of the official languages of Wales (alongside English, of course, which was the sole official language of the country for many centuries), and its use is strongly encouraged: for example, of people applying for certain jobs, those who speak both Welsh and English are preferred. (It must be added that there is no adult speaker of Welsh who does not speak fluent English, while most native English speakers have little or no command of Welsh at all.)

It must be emphasized that Welsh is not a dialect of English, but a Celtic language, which – though distantly related to English, both being Indo-European – is completely

[17] Unlike the name of the famous Scottish football club, which is pronounced ˈ**seltɪk**.

different, and it is not to be confused with **Welsh English**, the variant of English spoken in Wales. Let me give you a sentence in Welsh and its translation into English:

(58) *Dw i'n nabod rhywun sy'n siarad Cymraeg.*
'I know someone who speaks Welsh.'

I hope the example is convincing enough! Back to the point, this difference is especially important to emphasize because in English, the adjective *Welsh* is ambiguous: it can refer to the Celtic language, but it can also mean 'relating to Wales'. For example, the term *Welsh literature* is ambiguous: it can mean (i) 'literature written in Welsh', or (ii) 'literature written in Wales, in any language' (Dylan Thomas, for example, wrote in English, but his work is part of Welsh literature in the second sense)[18].

The other modern representative of the Brythonic branch of Celtic is **Breton**, spoken in the North-Western part of France called **Brittany** (in French – and, from this, in Hungarian - *Bretagne*). We have no exact statistics regarding the number of its speakers, because in France, no offical census is ever made regarding minority languages; according to reasonable estimates, the number of speakers is around 200,000 (but probably not more than 300,000 at best), the majority of speakers being over 60 years of age, which means that fewer and fewer young people speak it. This is a sharp decline: around 1950, the number of Breton speakers was around a million (so, at the time, it was spoken by more people than any other Celtic language)! Indeed, the UNESCO has listed it among the endangered languages of the world, and also, it is the only Celtic language which has no official status, not even regionally; this is because French is constitutionally declared to be the only official language of the French Republic.

As far as its origins are concerned, Breton, though spoken on the continent, is not a descendant of any ancient continental Celtic language. Instead, it (more precisely, its distant ancestor) was carried to Brittany from Britain, by Celtic immigrants who fled there from the Anglo-Saxon invaders during the 5th and 6th centuries AD – this is why it is closely related to Welsh. An interesting point to note is that the French name *Bretagne* derives from *Britannia*, the Latin name of Britain; the term *Great Britain* to refer to the largest of the British Isles has been in use since the Middle Ages to distinguish it from Brittany, i.e. 'Little Britain'. So the word *great* has nothing to do with a British sense of superiority or imperial might.

Finally, there is a third member of the Brythonic branch, **Cornish**, which is extinct (= dead). It used to be spoken in the South-West of England, in the area known as Cornwall. Some enthusiasts are attempting to revive it but without any real success.

2. **Goidelic**, the other branch of Celtic, is represented by two living languages. The first of them is **Irish Gaelic**, spoken in Ireland alongside English. It is mostly referred to simply as *Irish*, but (as in the case of *Welsh*) this is ambiguous, and again, it is not to be confused with **Irish English**. According to the latest (2011) census, the number of native or fluent Irish speakers (who use the language outside school on a daily basis) is about 80,000, but more than a million people claimed to use it to some extent; this, however, is misleading, because most of these people have actually very little command of Irish. This high figure is probably due to the fact that Irish is compulsorily taught at school in the Republic of Ireland, but very few people whose native language is English learn it really well. Irish is – although clearly a minority language – is the first official language of the Republic of Ireland according to the consitution (English being the second), and as such, the only Celtic language to be among the official languages of the European Union. In everday practice, however, it is used by a minority. In Northern Ireland (which, of course, is part of the UK), it enjoys regional

[18] In Welsh, the two senses are distinguished: the word *Cymraeg* is used in sense (i), but *Cymreig* in sense (ii).

status, although there, too, it is spoken by a minority of the population. Outside Europe, Irish speakers are found in relatively significant numbers in North America.

The other member of the Goidelic branch is **Scottish Gaelic**, spoken in Scotland by about 60,000 people. Geographically, it is basically restricted to the islands off the North-Western coast of Scotland, especially the Outer Hebrides, as well as parts of the Highlands; in the South and South-East of the country, it is practically not spoken. In spite of the small number of speakers, the language enjoys official status in Scotland, at least in the government of the country. There are small communities of Scottish Gaelic speakers in North America, but their number is not significant. Historically, Scottish Gaelic is the descendant of the language of immigrants from Ireland, who settled in Scotland sometime during the 4th and 5th centuries AD. This explains why Scottish and Irish Gaelic are quite similar, being, however, quite different from the Brythonic languages. As in the case of Welsh and Irish, Scottish is not to be confused with either Scottish English or Scots, which are variants of English.

Finally, the original language of the Isle of Man, called **Manx**, the third member of the Goidelic branch, is now extinct.

4.2.3 Italic

This branch has received its name from Italy, where it originates from. Today, it basically means the Romance languages, because the chief ancient representative of this branch is Latin (of which the modern Romance languages are descendants). Latin was originally only one of the Italic languages: the one spoken in Rome and a relativelly small area around it, known as Latium. With the rise of Rome and its expansion to the whole of Italy, and later to the vast area we call the Roman Empire, the other Italic languages were gradually displaced by Latin, becoming extinct (these include Oscan and Umbrian, for example). Therefore, the terms *Italic* and *Romance* are not synonyms historically speaking, in spite of the fact that all modern Italic languages are Romance ones.

Latin is first attested in the 5th century BC, and from the 3rd century BC onwards a vast amount of literature was produced in it. By the 1st century BC, a standard written form of Latin had come into existence. This form, called **Classical Latin**, was the norm which educated people used in writing and formal speech (such as public speeches). Indeed, this is the form of Latin taught in schools up to the present. The influence of Latin on the vocabulary of European languages can hardly be exaggerated; English has thousands of words of Latin origin, but even Hungarian has at least several hundred. Latin was in use as the language of science up to modern times, and up to the middle of the 20th century, it was used as the language of Roman Catholic services.

Classical Latin, however, was never used by ordinary people – in fact, not even by educated people in informal situations. Even though Cicero, for example, is one of the greatest authors in Classical Latin, he did not use it when talking to his friends over a bottle of wine. The spoken variety of Latin is called **Vulgar Latin**, from the Latin word *vulgaris*, meaning ‘popular, belonging to the people’; it has no negative connotations, therefore, contrary to the generally negative meaning of the word *vulgar* today. It must be emphasized that (as every spoken language), Vulgar Latin was not uniform: it showed considerable dialectal and social variation. The modern Romance languages are descendants of Vulgar Latin rather than Classical Latin; indeed, some of the differences that exist between modern Romance languages go back to ancient differences that existed between different dialects of Vulgar Latin. Let us now look at the Romance languages one by one, starting from the West.

1. **Portuguese**, the national language of Portugal is the first one. Since Portugal was one of the first European countries to colonize overseas territories, it is now spoken on several

continents. The largest country where Portuguese is the official language is Brazil, being not only the largest country in Latin America, but also the largest Portuguese-speaking country in the world. Portuguese is also spoken in parts of Africa and Asia.

2. **Galician**, spoken in the North-Western corner of Spain, north of Portugal. It is very close to Portuguese, especially to its neighbouring Northern dialects; so close indeed that we can regard it as a dialect of Portuguese. However, as it has been spoken in Spain, rather than Portugal, for many centuries, and it also has its own standard and official form, it is regarded as a separate language (but note that this is due to external, i.e. political and cultural reasons). It enjoys official status in the area where it is spoken (alongside Spanish, of course).

3. **Spanish**, the main language of Spain, which is also the nationwide official language. Because modern standard Spanish originates from the dialect of Castile, the province where Madrid is also found, it is frequently referred to as **Castilian** (a term especially popular in Latin America, where people speak Castilian, but they themselves are not Spanish). Spanish, like Portuguese, was among the first languages to be exported overseas – so successfully indeed that most of Latin America is now Spanish-speaking. It is interesting to note for students of English that, due to massive immigration to the USA from Latin America (especially Mexico), now a significant Spanish-speaking minority lives in the US, and in some places, Spanish speakers constitute the majority! The Spanish-speaking population of the US increases to such an extent that in a couple of decades, the country may become virtually bilingual.

4. **Catalan**, spoken in the East of Spain, chiefly in Catalonia and Valencia, but it is also the native language of the Balearic Islands (also belonging to Spain). In these areas, it is an official language. It is also an official language in the tiny country of Andorra, and it is spoken, furthermore, in the South-West of France (where, just like Breton, it has no official status). Although a minority language in the countries where it is spoken (except tiny Andorra), it is by no means a "small" language as far as the number of speakers is concerned, which is estimated to be close to ten million, as much as the entire population of Hungary.

There is a fourth language spoken in Spain, more precisely, in the north: **Basque**. This language, however, is not Romance – not even Indo-European. In fact, it has no known relatives at all! Attempts have been made to relate it to several different languages (I am quite sure Hungarian must be among them), but without any real success.

5. **Occitan**, historically closely related to Catalan, is also known as **Provençal** (although strictly speaking, the two terms are originally not equivalent). It is spoken in Southern France. The fate of Occitan clearly shows the aggressive, centrally forced expansion of French to the whole of France. In the Middle Ages, Occitan was spoken all over the southern part of France, and it was a highly prestigious language with a rich literary tradition (it was the language of the *troubadours*, for example); indeed, even the Italian poet Dante quotes entire lines in Occitan (!) in his Divine Comedy. By today, Occitan has become a minority language with (of course) no official status, chiefly spoken in rural areas, mostly by old people, and clearly doomed to die out.

6. **French** is spoken in several countries of Europe: France, the southern part of Belgium known as Vallonia, as well as in western parts of Switzerland. Its first documentation dates from the middle of the 9th century, making it the first Romance language to be written (at least no earlier text survives). It is no wonder that writing in the Romance languages starts quite late: for a long time, Latin continued to be used in written texts. One reason for this is the high prestige of Latin and its long written tradition, but there is another reason, too. Notably, language changes gradually and quite slowly. For some time after the collapse of the West Roman Empire, people could still understand Latin, and it was only when the local spoken language had become very different from Latin (so different that it was no longer intelligible to the average person) that sometimes, it was used in writing.

Back to French, a point of particular interest for the student of English is that it (more precisely its Norman dialect) was exported to England after the Norman Conquest of 1066, making it the language of high social classes for some centuries, so at the time, England was one of the places where French was spoken! This issue will be discussed in detail in the chapter on Middle English. Another interesting point is that French is also spoken in Canada, in the province of Quebec, where it is the native language of the majority, who have a strong sense of national pride, and the question of Quebec becoming independent from the (English-speaking) rest of Canada is constantly on the agenda. French is the official language in Quebec, and, on a federal level, one of the official languages of Canada, too. The reason for French being spoken in Canada is that originally, much of North America was colonized by France, but it gradually lost its territories there. Most of North America became English-speaking, but Quebec remained French-speaking (although under British rule for a long time). The French language was also carried to other parts of the world via colonization, especially to North-Western Africa, where it is still an official language in many countries (e.g. Algeria). It must be noted, though, that French has never replaced the native languages of these countries, so they have not become truly French-speaking. The reason for this that there has never been massive emigration to these countries from France. Educated people still learn French, and (as I said) it is an official language, but it is not widely spoken in everday life. May I note that the situation is quite similar in the former British colonies in Africa regarding the use of English, with the notable exception of the Republic of South Africa (where English is spoken by a significant part of the population as a native language).

7. **Italian**, spoken in Italy (where it is the national language) and in the south of Switzerland. Spoken Italian, in fact, shows extreme dialectal variation, and several dialects are so different that they are mutually unintelligible – so their speakers need to use standard Italian to converse with each other. Standard Italian has had a long tradition, going back to the Middle Ages, when literature in Italian came into existence, chiefly thanks to the three great poets, Dante, Boccaccio and Petrarch, during the 13th and 14th centuries. Since they wrote in their own dialect, which was that of Florence, it was this dialect which became the basis for standard Italian later on (and which is why, up to the present day, this dialect is closest to the standard). Since Italy was not one of the great colonizers, the language has not spread overseas to a significant extent, but Italian emigrants to some countries have carried their language with them; as a result, Italian is spoken by a significant number of people in the US, for instance.

8. **Rhaeto-Romance**, which is not really a language but rather an artificial term used for a group of closely related Romance languages spoken in the South-East of Switzerland and the North-East of Italy. The Rhaeto-Romance language spoken in Switzerland, called Romansh, is – alongside French, Italian, and, of course, German – one of the national languages of Switzerland.

9. **Romanian** is the official language of Romania and the Republic of Moldova. It is also spoken by a small number of speakers in Hungary, the Ukraine, and Serbia. Interestingly, there are dialects of Romanian spoken in Macedonia and even as far away as the Istrian peninsula in Croatia. The main body of Romanian-speaking territories is geographically quite separated from the rest of Romance. There are competing theories about the origins of Romanian: the official Romanian view is that the language is a descendant of the Latin of the province of Dacia, but it is much more likely that the language originates from the Romance-speaking areas in the South-West of the Balkans, close to the Adriatic Sea, and it was this dialect which, through migration, was carried to the present-day Romanian-speaking lands. An important argument for this view is that Romanian shows a close degree of relatedness to **Dalmatian**, a Romance language formerly spoken in Dalmatia (roughly, the middle and the southern part of the Croatian coastal area), which is now extinct, but there are written sources.

On the other hand, contrary to popular misconception, often advertized by Hungarian extremists, Romanian is unquestionably a Romance language, deriving from Latin, rather than Latinized Slavonic or whatever else. The myth of Romanian being Slavonic probably derives from the fact that the language was written in the Cyrillic script before 1859. This is indeed unique among Romance languages, but writing is a matter of cultural tradition which often does not have anything to do with the origins of a language – indeed, most European languages are written with Roman letters, including Hungarian, which is not only not Romance, but not even Indo-European! The fact that Hungarian uses the Roman alphabet does not make it Latin or Romance, and neither is Romanian Slavonic. This discussion leads us to the next large branch of Indo-European.

4.2.4 Slavonic

This branch is also called **Slavic**, especially in North America. Its members cover most of the eastern half of Europe, but one of them, Russian, is spoken in huge territories in Asia, extending to the Pacific Ocean. Since we mentioned the question of writing in the previous section, I would like to note that some Slavonic languages use the Roman alphabet, while some use the Cyrillic one. Slavonic languages are traditionally divided into three groups: West, South and East Slavonic.

1. **West Slavonic** includes Czech, Slovak, Polish and Sorbian. **Czech** (pronounce exactly like the word *check*) is spoken chiefly in the Czech Republic. It is closely related to **Slovak**, - so closely indeed that speakers of these two languages can understand each other very well. Indeed, in the 19th century as well as the early 20th, there were (sometimes fierce) arguments whether Slovak should be regarded as an independent language or a dialect of Czech. Since Slovak was spoken in Hungary (before the Treaty of Trianon in 1920), however, it came to be regarded as an independent language. Slovak is spoken today chiefly in Slovakia, but there are many speakers living in the Czech Republic, too, most of whom moved there during the time when Czechs and Slovaks lived in the common state of Czechoslovakia. Scattered Slovak communities are also found in Hungary.

Polish is the largest of the West Slavonic languages in terms of the number of speakers. It is spoken mainly in Poland, but – interestingly for students of English – there has been massive Polish immigration to the UK, especially London, since Poland joined the European Union in 2004; as a result, there is now a significant Polish-speaking community in the UK, amounting to tens of thousands of people. Several Polish-speaking immigrant communities are also found in the USA.

Sorbian – not to be confused with Serbian! – is a language spoken in the eastern parts of Germany, close to Poland, mainly around the town of Lausitz. It is the only West Slavonic language which is not an official one and which is a minority language in the country where it is spoken.

2. **South Slavonic** includes several languages. Starting from the West, **Slovene** (also called **Slovenian**) is spoken mainly in Slovenia, but there are autochtonous Slovene communities in Italy (around Trieste), Austria (in southern Carinthia) as well as in Hungary (the area traditionally called *Vendvidék* in Hungarian, the small triangular area south of Szentgotthárd; *vend* is an old word for Slovenian). Slovene has a particularly interesting grammatical feature from an Indo-European viewpoint, which will be mentioned in Section 4.3. below.

The next language is **Croatian** – but here, we encounter a problem of a political nature. Croatian is closely related to **Serbian**. In fact, this is an understatement: the differences between the two are no bigger than between standard British and American English! Linguistically speaking, therefore, we might as well consider them variants of one language (think of Galician and Portuguese), and indeed, they were regarded as such during the time when Croats and Serbs lived in the state of Yugoslavia. They were together called **Serbo-Croat** officially. It must be emphasized, though, that in everday life, people did not use this term, Croats using the term Croatian, Serbs using Serbian (as they have always done). After the violent breakup of Yugoslavia, the term Serbo-Croat ceased to be used officially, too. The point is that whether something is regarded as one language or two (or more), is a matter of cultural and political tradition: whether Serbians and Croatians wish to consider their language the same or not is up to them, and they are free to decide. The more difficult problem is that basically the same language is spoken by Bosnians (the Muslim population of Bosnia and Hertzegovina) as well as by Montenegrins; since they are neither Serbs nor Croats, the problem has arisen what their language should be called. At any rate, a detailed discussion of this would be beyond the scope of the present book. As a point of interest, however, I would like to add that Croatians (as well as Bosnians) use the Roman alphabet; as opposed to this, Serbians (and Montenegrins) have traditionally used the Cyrillic script, although both in Serbia and Montenegro, the Roman script is also used. Therefore, the same text in Serbian can be written in two ways, although pronounced the same! Since most students of English are nowadays not familiar with the Cyrillic script, I would – as a point of interest – like to give an example of a sentence in Serbian, written in both ways:

(59) Cyrillic: Молим вас, где је станица?
Roman: Molim vas, gde je stanica?
In English: ‘Where is the station, please?’

Let us now move on to the rest of the South Slavonic branch, including two closely related languages: **Macedonian** and **Bulgarian**, spoken in Macedonia and Bulgaria, respectively. Macedonian was officially recognized as an independent language as late as the middle of the 20th century by the then Yugoslav government (Macedonia was part of Yugoslavia) in order to emphasize its difference from Bulgarian, because Macedonia was also claimed by Bulgaria (indeed, the official Bulgarian view is still that Macedonian is but a dialect of Bulgarian). This again illustrates the political and cultural complications involving linguistic identity, but I do not wish to discuss the issue in detail. Both languages use the Cyrillic script, and it is now high time we talked a bit about the use of the Cyrillic script.

As noted, it is used by Serbs, Montenegrins, Macedonians, Bulgarians – and also by all East Slavs. What is common to these peoples is that they all belong to the Eastern (Greek) branch of Christianity, as opposed to the rest of the Slavs, who are Roman Catholic. The Cyrillic script itself is named after St. Cyril, who – together with his brother, St. Methodius – created the first Slavonic script in the 9th century AD. In fact, the alphabet created by them was not Cyrillic, but a rather different one, called **Glagolitic**; it is a very strange and unique alphabet, whose origins are mixed: some letters seem to originate from the Roman script, some from Greek, some from Hebrew, but several of them have no clarified source. Glagolitic is no longer in official use anywhere, but you may come across Glagolitic inscriptions in churches in Croatia (where it remained in use – at least by the Church – relatively long). The Cyrillic alphabet itself is basically a modified version of the Greek alphabet, which was not used before the 10th century. The reason why I mention this at this point is that St. Cyril and St. Methodius spoke Old Bulgaro-Macedonian, and it was this variety of early Slavonic in which the first Slavonic texts were written; it is also known as **Old Church Slavonic**.

3. **East Slavonic** includes three languages. The one with the highest number of speakers is **Russian**, which is – in terms of the number of speakers as well as geographically – the "largest" of all Slavonic languages in general. Being the official language of the Russian Empire, then of the Soviet Union, it is spoken in a number of countries outside present-day Russia as well: the eastern part of the Ukraine, for example, is mostly Russian-speaking, but there is also a significant Russian minority in the Baltic states of Estonia, Latvia and Lithuania, to mention but a few examples. The official status of Russian outside Russia (I mean the former Soviet republics) is a complex and controversial issue, due to political reasons.

Belarusian (also known as **Belorussian**) is the national language of Belarus – although it may actually be spoken by a minority, since much of the country is Russian-speaking. The third language in the East Slavonic group is **Ukrainian**, spoken in the Ukraine; its speakers are concentrated in the western half of the country (the East, as mentioned, being chiefly Russian-speaking).

This concludes our discussion of Slavonic; let us now turn to the remaining branches of Indo-European spoken in Europe, then move on to Asia.

4.2.5 Baltic

This branch is represented by two living languages. The first of them is **Lithuanian**, the official language of Lithuania, and the other one is **Latvian**, spoken in Latvia. In both countries (as mentioned above) there is a significant Russian minority, too. It must be strongly emphasized that – although Estonia is, geographically speaking, a Baltic country – the Estonian language itself is not a member of the Baltic branch of Indo-European: it is Finno-Ugric (so it is not Indo-European at all)!

There is an extinct member of the Baltic branch, too: **Old Prussian**, which was spoken in the area known as Prussia. It had become extinct due to various reasons by the early 18th century, and the area became chiefly German-speaking. The interesting point is that the name of the land remained in use, and, since the Kingdom of Prussia played a leading role in German unification in the 19th century, the term Prussian became strongly associated with Germans and Germany. Nonetheless, Old Prussian is not a variety of German! After World War II, the territory was annexed partly to Poland, partly to the Soviet Union, and Germans were forcibly exiled, so that Prussia today is really but a historical term.

4.2.6 Albanian

This branch contains one language only: Albanian. This means that the language is Indo-European, but it has no close relatives (it is a "single child", as it were). It is spoken in Albania as well as Kosovo, where Albanian speakers constitute an overwhelming majority. Kosovo was an autonomous area of Serbia, but it declared independence in 2008. Nonetheless, its independence is practically only half-existent, since the majority of the world's countries have not recognized it as an independent state up to the present. Albanian is also spoken by a significant minority in Macedonia, and by scattered communities throughout the former Yugoslav republics.

4.2.7 Hellenic

This branch contains one member only: **Greek**, which, however, is one of the most influential Indo-European languages, due to its ancient written tradition. On the one hand, alphabetic writing (which means that letters represent phonemes rather than entire words or syllables) was invented in Ancient Greece. The Greek alphabet was also adopted – although in a modified form – by the Romans, so the Roman alphabet is Greek in origin (as we mentioned earlier, so is the Cyrillic script). Second, the flourishing literary culture of Ancient Greece had a decisive influence on Roman literature – as well as the Latin language, which adopted many Greek words. Due to the influence and high prestige of Latin (and Roman culture) throughout the history of Europe, the influence of Greek has remained substantial up to the present; needless to say, ancient Greek literature – and the language – have themselves remained in the focus of attention. Hundreds of words used in European languages are of Greek origin, including a number of technical and scientific terms (e.g. *atom, psyche, hydrogen, biology*) but some common ones, too (such as *museum*, or, indeed, *music*)[19].

In fact, another reason why Greek has been so influential is that it is the language in which the books of the New Testament were written. This requires a brief comment, after all, you may wonder why the New Testament was written in Greek rather than Latin (the official language of the Roman Empire) or Aramaic (spoken by Jews at the time, the native language of Jesus), or possibly Hebrew (the language of the Old Testament). Now, Aramaic and Hebrew were understood by Jewish people only, and Christianity was from the outset an "international" religion to be spread in the world. Latin was never widely spoken in the Eastern Mediterranean: it was used as the language of administration and the military forces, and Latin was certainly understood by educated people, but the common folk never became Latin-speaking, and most people did not understand Latin at all. As opposed to this, Greek had been in use in the Eastern Mediterranean for a long time as a common mediating language (called **lingua franca**) among the different peoples of the area[20]. This is chiefly because of the leading role that Greece played in commerce; at any rate, Greek was by far the most widely understood "universal" language of the Eastern Mediterranean, which also happened to have had a long and rich literary tradition, so it must have seemed to be the best choice.

Before about the 3rd century BC, literature in Greek was written in various dialects, i.e. there was no single standard. Starting at this time, however, the dialect of Athens (by then the leading city-state among all Greek ones), called **Attic**, gradually evolved into a common standard named *koine* (indeed, a word meaning 'common'). It was, in fact, this Attic-based variety which spread all over the Eastern Mediterranean as a lingua franca, and, accordingly, it is the variety of Greek in which the New Testament was written, and indeed, the one from which Modern Greek has evolved.

At present, Greek is spoken chiefly in Greece and part of Cyprus, but there are significant emigrant Greek communities in the UK as well as in the USA.

4.2.8 Armenian

Yet another branch which is represented by one language: Armenian. It is spoken in Armenia and in neighbouring countries. Historically, Armenia was also spoken in what is now Eastern

[19] Greeks are immensely proud of this: think of the film *My big fat Greek wedding*, in which one of the characters claims every word to be of Greek origin, a source of constant amusement throughout the film.

[20] This can be likened to the present-day international status of English, which is today's global lingua franca: if people of different nationalities come together, they are most likely to use English.

Turkey, but the Armenian population of that area was exterminated by the Ottoman authorities during World War I (chiefly in 1915). This is known as the Armenian Genocide (or Armenian Holocaust), and the number of victims is estimated to have been between 1 and 1.5 million. (May I add that Turkish authorities deny this genocide up to the present.)

Armenian has a special alphabet of its own, and it has a long and rich literary tradition; furthermore, Armenia was the first country to adopt Christianity as a state religion, at the very beginning of the 4th century AD.

4.2.9 Indo-Iranian

Indo-Iranian is divided into two branches.

1. **Indic**, spoken chiefly on the Indian subcontinent. This branch includes a number of languages, of which (regarding the number of speakers) the largest ones are **Hindustani**, which is called **Hindi** in India and **Urdu** in Pakistan, and **Bengali**; there are many other, but I will not list all of them here. **Romany**, spoken by many Gypsies in Europe including Hungary, also belongs to the Indic branch. It must be noted, though, that not all Gypsies speak Romany – indeed, most of those who live in Hungary speak Hungarian as a native language.

Historically, Indic has had a long literary tradition, going back to more than 2,500 years. The most ancient form of written Indic is known as **Vedic**, the language of the Vedas, the founding texts of Buddhism. In the 4th century BC, the Indian grammarian Panini standardized the language, creating what is known as **Classical Sanskrit**. As in the case of Latin, however, the spoken language was different, and indeed, it had several varieties known as **Prakrits**. The modern Indic languages are descendants of ancient Prakrits, just like the Romance languages are descendants of local varieties of Vulgar Latin.

2. The other branch within Indo-Iranian is **Iranian**. The best known representative of the group is **Persian**, the national language of Iran. Be careful: although Iran is an Islamic state and they use the Arabic script, Persian is an Indo-European language! The other two major languages belonging here are **Pashto**, spoken in Afghanistan, and **Kurdish**, spoken in several countries including Iraq, Iran, Armenia and Turkey. Ancient Hungarians – on their way to Central Europe – also came into contact with Iranian-speaking peoples, from which several Hungarian words were borrowed (e.g. *vásár, asszony, híd*, etc.).

A final interesting point: Indo-Iranian languages are spoken by an estimated 1.5 billion people – amounting to half of the total number of Indo-European speakers.

4.2.10 Other

There are two other branches of Indo-European, spoken in ancient times. One of them is **Anatolian**, the chief representative of which is **Hittite**; the other one is **Tocharian**. Neither of them have any living descendants, so I am not going to discuss them in detail.

This concludes our discussion of Indo-European languages. Let us now turn our attention to the parent language of this huge family, Proto-Indo-European.

4.3 Proto-Indo-European

As mentioned earlier, the parent language of the Indo-European family, called Proto-Indo-European (from now abbreviated PIE), was spoken about 6,000 years ago, presumably in

Central-Eastern Europe (although opinions vary – some argue for Anatolia, for example). Let us now sum up the main features of this language.

It must be emphasized that PIE is not attested: no documentary evidence is available. What we know about it is the result of reconstruction: using historical linguistic methods, and comparing the Indo-European languages (especially the very early ones, which are obviously closer to the parent language than the modern languages), we can reconstruct its chief properties. Needless to say, there are debates among scholars regarding the reconstruction of PIE; I will present a fairly consensual view here, without going into details.

4.3.1 The sounds of PIE

1. The vowel system of PIE is a fairly average one:

(60) Short vowels: **i u e o a ə**
Long vowels: **iː uː eː oː aː**

As you can see, the schwa only occurs as short, otherwise the short and long vowels are arranged neatly into pairs. Diphthongs are often reconstructed, too, but it is a debated issue and I will not use them here.

2. The consonant system of PIE, however, is highly interesting, specifically, the system of obstruents, which contains a number of stops (plosives) but only one fricative. Look at (61) below:

(61)

Place of articulation →	Labial	Dental	Alveolar	Palatal	Velar	Labio-velar
A. Sonorants						
Nasals	**m**	**n**				
Liquids			**l, r**			
Glides	**w**			**j**		
B. Obstruents						
1. Fricative			**s**			
2. Stops						
Voiceless	**p**	**t**			**k**	**k^{w}**
Voiced	**b**	**d**			**g**	**g^{w}**
Voiced aspirated	**b^{h}**	**d^{h}**			**g^{h}**	**g^{hw}**

For an illustration of the pronunciation of unusual symbols, see the notes on page 4.

As mentioned above, there is a rich set of stops, but there is only one fricative, **s** (which is voiceless). Indeed, a wholesale transformation of this system is a characteristic feature of Germanic, which will be discussed in the next chapter (indeed, all Indo-European languages increase the number of fricatives to some degree).

4.3.2 The reconstruction of PIE morphemes: An illustration

With the help of scientific methods, we can reconstruct the shape of morphemes in PIE, both roots and affixes. Look at the example below, showing the root morpheme meaning 'to bear, to carry' in some Indo-European languages.

(62)

English:	*bear*, cf. also Old English *ber*
Latin:	*fer*
Ancient Greek:	*p^{h}er*
Czech:	*ber*
Sanskrit:	*b^{h}ar*

As you can see, the forms of this root, although not identical, are still quite similar in the various languages: they all begin with some labial consonant, end in **r**, and – with the exception of Sanskrit – they have the short vowel **e** (in PdE, you have a long vowel, but this is a relatively recent development: Old English still had a short **e**). With the help of the reconstructive method, linguists have assumed a PIE form **b^{h}er-*; the asterisk (star) before the form shows that the form is reconstructed, rather than attested. From this proto-form, the forms of the daughter languages can be derived via regular sound changes. In Sanskrit, for example, short ***e** regularly becomes **a**, but it remains in the other languages. Note that the initial consonant (= ***b^h**) remains unchanged in Sanskrit only: in Slavonic (including Czech) as well as in Germanic, it loses its aspiration, becoming simple **b**, while in Greek, it remains aspirated but becomes voiceless (= **p^h**); finally, in Latin, it also turns into a voiceless fricative (= **f**). Note that the PIE form itself does not remain unchanged in any of the daughter languages.

Needless to say, the above example is meant as an illustration only; if you are interested in the details of linguistic reconstruction, please check the Suggested Reading section at the end of this chapter.

4.3.3 The morphology of PIE

4.3.3.1 Nouns, adjectives and pronouns

PIE was a very richly inflected language, using mostly affixes to express grammatical functions such as number, case, gender, etc. Nouns and adjectives, as well as certain pronouns, distinguished **three grammatical genders: Masculine, Feminine and Neuter**. This threefold gender system is still found in German as well as in the Slavonic languages; it also existed in Old English, but it was lost by the Middle English period. In Latin, we have the same system, too, although the modern Romance languages (with the possible exception of Romanian) have only two genders (Masculine and Feminine) left – if you happen to speak French, Italian, Spanish, etc., the situation will be familiar.

It is extremely important to emphasize that grammatical gender has nothing to do natural gender, i.e. sex. As the name suggests, it is a purely grammatical category. In practice, it means that words belonging to different genders are inflected differently, and they also take different forms of articles, for example. In German, for instance, Masculine nouns take the form *der* of the definite article (e.g. *der Tisch* 'the table'), Feminines take *die* (e.g. *die Sprache* 'the language'), while Neuters take *das* (e.g. *das Schiff* 'the ship'). Note that the

meaning of these nouns has nothing to do with natural gender: a table is not a male being, neither is language a female thing, for example. In Slavonic languages, gender distinctions are generally shown by the ending a noun takes: in Slovene, for instance, nouns ending in a consonant (in the Nominative Singular, that is, the dictionary form of the noun) are mostly Masculine (e.g. *brat* 'brother', *most* 'bridge'), those in *–a* are typically Feminine (e.g. *roka* 'hand', *mačka* 'cat'), while Neuters typically end in *–o* or *–e* (e.g. *mesto* 'town', *morje* 'sea').

The gender of a noun is fixed: a given noun belongs to a particular gender. Adjectives and pronouns, on the other hand, have varying forms, according to the gender of the noun they stand together with. In Slovene, for instance, you have the adjective meaning 'beautiful' in varying forms: *lep most* 'beautiful bridge' (Masculine), *lepa mačka* 'beautiful cat' (Feminine), and *lepo mesto* 'beautiful city'. The sitation was somewhat similar in Old English, where some endings were characteristic of a certain gender (or genders); we will see some examples in Chapter 5.

An interesting point about pronouns is that personal pronouns did not distinguish gender in the first and second persons. (They did in the third persons, but originally, these pronouns derive form earlier demonstratives.) Now, in Present-day English, we have no grammatical gender any longer, but a trace of this categorization still remains: in the third person singular, we have three personal pronouns: *he*, *she*, and *it*. In Old English, these originally differed in grammatical gender; after grammatical gender had been lost, these three pronouns still remained, but today, they are used according to natural gender (that is, *he* refers to male persons, *she* to female persons, and *it* is used to refer to sexless things (or to beings, such as animals, whose sex is irrelevant). Now if you compare this threefold distinction to other personal pronouns, none of them expresses a difference in gender: *I*, *you*, *we*, *they* are all used to refer to male or female or sexless beings or things. (In fact, the third person plural did originally have three different forms, too, but this distinction was lost very early in the history of English.)[21]

The second important category, present through the whole morphology of PIE, is a threefold distinction in number. Most modern Indo-European languages distinguish two numbers: **Singular** and **Plural**. In PIE, there was a third number: **Dual**, used to refer to two things or persons. A Slavonic language used above as an illustration for gender, Slovene, is an exceptional language among the modern Indo-European ones inasmuch as it still has the Dual number. Remember the word *mesto* 'town'? Well, in English, if you talk about more than one town, you use the plural: *towns*. In Slovene, however, there are two corresponding forms: *mesti* 'two towns' and *mesta* 'more than two towns'. Note that the form *mesti* already includes the information that you talk about two (and not more) towns, without actually using the numeral 'two' (you can say *dve mesti* 'two towns', but only if you want to put emphasis on the numeral: the form *mesti* alone means the same). In Old English, the Dual number survives with some personal pronouns, but otherwise it had been lost rather early. You still have relics of the Dual, though: the concept 'all of the two', for example, is expressed by a separate form, viz. *both*, originally a Dual form (cf. *all countries* vs. *both countries*). Similarly, *either* and *neither* are used to express 'any of the two' or 'none of the two' (cf. *either/neither country* vs. *any/none of the countries*). This does not mean, however, that the Dual as a systematic category still exists in English: as I said, these are but relic forms.

Nouns, adjectives and pronouns also distinguished **case**. If you recall from Chapter 2, in Modern English, personal pronouns (more precisely, some of them) distinguish Nominative, Accusative, and Genitive. Nouns, on the other hand, have only two inflected

[21] In some cases, *he* or *she* is used to refer to non-living (sexless) things: for example, sailors typically refer to their ship as *she*. This, however, is an instance of personification, and it has nothing to do with grammatical gender; neither is it a survival of Old English gender (the noun *ship*, for example, was a Neuter one in Old English).

forms: Nominative and Genitive (e.g. *boy* and *boy's*): the Accusative is the same as the Nominative (so *boy* can be both, cf. *The boy arrived* vs. *I found the boy*). As for PIE, we can reconstruct as many as 8 cases. I list these now, together with their main functions[22].

1. The **Nominative** – as in English – is the form of subjects, as shown in English by the pronoun *I* in *I work a lot*.

2. The **Accusative** is primarily the case expressing direct object (as in *I found him*). It also occurs after certain prepositions – we say that certain prepositions **govern** the Accusative case. In Modern English, all prepositions govern the Accusative, but in PIE – indeed, still in Old English – prepositions varied according to the case they governed, so not all of them took the Accusative: some governed the Genitive, for instance, or some other case form. The Nominative never stood with prepositions.

3. The primary function of the **Genitive** was to express possession, as in English, but it could also express other functions. The most important of these is the so-called **partitive** function, expressing a certain amount of something. Indeed, in some cases, even Modern English uses an *of*-genitive for the same purpose, cf. *a bottle **of** milk, a lot **of** people* (compare Hungarian *egy üveg tej, sok ember*), but note that the *of*-genitive structure appears as late as in Middle English, so it is not a direct continuation of the PIE Genitive: the point is that functionally, it behaves similarly. What is interesting to note is that the *s*-genitive (e.g. *boy's*) does derive from an ancient PIE Genitive suffix *–s*, compare Latin *pater* 'father' and *patris* 'father's', for instance.

4. The **Dative** is the case expressing indirect object, as in *I gave him a book*. In Modern English, prepositional phrases with *to* are also used in the same function, as in *I gave the book to him*. I would like to emphasize that English no longer has a Dative case: personal pronouns which distinguish the Nominative from the Accusative (e.g. *he* vs. *him*) use the Accusative in an indirect object function. In Old English (if you recall Chapter 1), there was still a different Dative form. The Dative also frequently occurs after certain prepositions.

5. The **Instrumental** is a case expressing means/instrument (= by/with what, as in *I travel by train*). In PIE, there was a separate case form, and this is still what we find in some Slavonic languages. In Czech, for example, 'train' is *vlak*, and its Instrumental form is *vlakem*, so 'I travel by train' is *Cestuji vlakem* (the verb form *cestuji* means, of course, *I travel*, like *utazom* in Hungarian, so the pronoun meaning *I* can be omitted). Note that the form *vlakem* expresses means/instrument alone, without a preposition. The Instrumental only survives marginally by Old English times, and by Middle English it is completely lost. There is one interesting relic form, though: the interrogative pronoun *why*, which was originally the Instrumental form of *what*, meaning 'by/with what'. During the centuries, however, its meaning gradually shifted to 'by what reason', i.e. 'why', becoming an independent word, no longer felt to be related to *what*, though historically, it is. The Instrumental, too, was found after certain prepositions.

6. The **Ablative** case was used to express source/origin, corresponding (roughly) to English prepositional phrases with *from*, as in *He came from London*. The Ablative had been lost very early in Germanic, and practically no trace of it is found in English. As other, non-Nominative cases, the Ablative could also occur with certain prepositions.

7. The **Locative** is primarily the case expressing location, corresponding to *at/in/on* in English, but it was also used together with some prepositions. This case does not survive into Germanic, either.

8. Finally, the **Vocative** was the form used to address someone, so it would be used in a sentence like *Father, please give me twenty dollars*. The famous sentence uttered by Julius Caesar when he was stabbed to death, i.e. *Et tu, mi fili, Brute* 'Even you, my son, Brutus' is a

[22] The Hungarian names of the cases (from 1 to 8) are as follows: 1. alanyeset, 2. tárgyeset, 3. birtokos eset, 4. részeshatározó eset, 5. eszközhatározó eset, 6. távolító eset, 7. helyhatározó eset, 8. megszólító eset.

well-known example from Latin, where the Vocative form *Brute* is used instead of the Nominative *Brutus*. The Vocative is found in some early Germanic languages (e.g. Gothic), but only marginally, and by Old English times, it practically disappears.

This concludes our discussion of PIE nouns, adjectives and pronouns. There would still be a lot to say – but that would be beyond the scope (and aims) of the present book. You are invited to check the Suggested Reading section for further material on the topic if you are interested.

4.3.3.2 Verbs

The system of verbs is highly complex in PIE from a morphological point of view, and I do not intend to cover the topic exhaustively. Instead, I restrict my attention to the most important points, and I also simplify a bit.

Firstly, just like nominals, verbs also distinguished **three numbers – Singular, Dual, and Plural**. In Old English, this distinction no longer exists, but verbs still have three numbers in Gothic, the earliest extensively documented Germanic language. For example, *slepa* (Sg., 'I sleep') differs from *slepos* (Dual, 'we two sleep') as well as *slepam* (Pl., 'we (= more than two of us) sleep'). Categories such as **case** and **gender**, which are nominal categories, were not expressed by verbs, of course. Just like personal pronouns, however, verbs also distinguished **person**, i.e. 1st, 2nd, and 3rd, in all three numbers.

Secondly, there were – as in all languages – characteristic verbal categories, specifically, **voice, aspect, tense** and **mood**.

1. There were two **voices**: **active** and **middle**. The active voice had the same function as in English, while the middle voice expressed reflexive as well as passive functions. The middle voice does not survive into English, although in the earliest periods of Germanic, some of its forms still remain, but mainly in a passive function. The details are not relevant for us, but I would like to note that the middle voice was originally expressed by affixation, and the Modern English way of forming a passive construction (= *be + Past Participle*, as in *the ball was found*) has nothing to do with the PIE middle voice: the English construction is a much later, independent development. Compare this to Gothic, where *nimand* means 'they take' while *nimaindau* is a passive form (= 'they are taken'), for instance.

2. Modern English distinguishes three aspects: **simple**, **perfect**, and **progressive/continuous**, compare *I work* vs. *I have worked* vs. *I am working*. (The "simple aspect" really means that it is neither perfect nor progressive.) In PIE, the situation was similar, though not identical. There was a **perfect aspect**, roughly corresponding to the perfect aspect of Modern English. The second one is called **imperfect**, meaning 'non-perfect', corresponding to the English progressive, but partly also to the simple: apart from progressive actions, it also expressed habitual events (as in *I go to work by bus every day*) or states which are generally true (as in *The Sun rises in the East*). The third aspect is called **aorist**[23], which expressed single, completed events – as in *John entered the room* or *The vase broke*. In most cases, then, an aorist form corresponds to a simple form in English, especially a simple past one.

3. Whether PIE originally had special forms to express different tenses or not is a debated issue. It is possible that tense distinctions were not made: instead, the different aspects served to express the difference between present, past, and future times. (Let me remind you that **tense** is a grammatical category: so, for example "future tense" does not

[23] Pronounce ˈ**eərɪst**.

mean the same as "future time"; if you say *The train arrives at 5 p.m.*, the sentence refers to future time, but *arrives* is grammatically speaking a present tense form!) The question remains unsettled, but it is not very important for us anyway, since the aspect/tense system was completely reorganized in Germanic.

4. There are four moods we can reconstruct for PIE: **Indicative, Imperative, Subjunctive** and **Optative**. The first two (= Hungarian *kijelentő mód* and *felszólító mód*) hardly need a detailed discussion. The Subjunctive (Hungarian *kötőmód*) was used to express uncertainty, doubt, or unreal situations, so it was also used in a conditional function. The Optative (Hungarian *óhajtó mód*) expressed wishes; in Germanic, it fell together with the Subjunctive – so in the Germanic languages, the distinction between the two is not made[24].

This concludes our brief overview of the chief morphological categories of PIE. Contrary to what you might expect, I am not discussing PIE syntax here. Firstly, there is relatively little that we know about PIE syntax, at least when we compare it to our considerable knowledge of PIE phonology and morphology. Secondly, the syntactic patterns of PIE have very little, if any, direct relevance for an understanding of the origins of Modern English syntax. The reason for this is that the syntactic structure of English underwent complex and extensive changes after the Old English period, so the older patterns mostly survive only as relics (if at all).

4.3.4 Ablaut

This phenomenon, found in PIE, and also surviving to varying degrees in the daughter languages, is on the boundary of phonology and morphology. The term Ablaut is a German word, but it is also commonly used in English, although the Greek-based coinage *Apophony* is sometimes found, too. It basically means a series of vowel alternations. Let us see what this means.

Remember the PIE root meaning 'to bear, to carry'. We have given the reconstructed form *b^her. The same root, however, may occur with other vowels, too, in the Indo-European languages. In English, for example, its past tense form is *bore*, its past participle is *born(e)*. (The last two are pronounced with the same vowel in RP: **bɔː** and **bɔːn**, but this is relatively recent: in EModE, they sounded different, viz. **boːr** and **bɔrn**; go back to Chapter 3 to check the RP developments. Indeed, the two forms still have different vowels in Scottish English, for example.) Now, have you ever guessed that the word *birth* was originally derived from *bear*? Yes, it was: *bear* has a meaning 'to give birth to a child' (after all, a pregnant woman does bear, i.e. carry, her child – compare also *kihord* in Hungarian), and *birth* was used in the sense of 'the act of bearing a child'. Note the different vowel! In fact, the noun *burden* (Hungarian *teher*), i.e. 'something heavy that you carry', also derives from *bear*. This variety of different vowels ultimately originates from an ancient, PIE, variation, known as Ablaut.

In essence, it is assumed that originally, all PIE roots had the basic vowel **e**, cf. the form *b^her. Due to some sound changes, however, this original vowel was replaced by other vowels due to some phonetic conditions (such as different stress), but it could even be dropped totally. Altogether, by late PIE, roots could have as many as five different vowels:

(63) **e** **eː** **o** **oː** **Ø** (= zero)

[24] A point of interest: it was the forms of the Optative which actually survived into Germanic, but primarily with a Subjunctive function.

These vowels, then, underwent their usual developments in the daughter languages, including English, so the Present-day English vowels are usually quite different from those in (63). In fact, *bear* shows quite complex developments, but take a simpler case: the verb *sing*, whose past form is *sang*, and the past participle is *sung*. The root is an ancient Indo-European one, which sounded something like **seng^{hw}-* in PIE. This is the basic form, from which *sing* has developed; in Germanic, PIE ***e** often becomes **i**. *Sang* derives from the PIE variant with **o**, i.e. **song^{hw}-*; in Germanic, PIE ***o** regularly becomes **a**, still found in the Middle English form (the modern **æ** is an EModE development from earlier **a**). *Sung* originates from **sng^{hw}-*, where the original **e** was dropped; the PIE form was pronounced with a syllabic **n**, which regularly developed into the combination **un** in Germanic, yielding *sung* – remember, the Present-day English vowel **ʌ** comes from EModE **ʊ**. (The development of the final consonant is not important for the present discussion.)

Ablaut is of relevance for an understanding of Modern English primarily because several irregular verbs owe their different vowels to it. Further examples include *swim/swam/swum, ride/rode/ridden, break/broke/broken*, to mention but a few. A more detailed description of such verbs will be provided in the next chapter.

You have also seen that Ablaut is also found in word-formation, cf. *birth* and *burden* above, both from *bear*. Such forms are less conspicuous than past or past participle forms, since their relatedness to each other is by today obscured due to several reasons: while anyone who speaks English will find it obvious that *sing* and *sang*, for instance, are related to each other, the relatedness of *bear* and *birth* is much less obvious – indeed, you need proper historical study to discover it. Let me, therefore, mention just one or two further word pairs whose members show a vowel difference originating ultimately from PIE Ablaut. The verb *set*, for example, originally meaning 'to make sg/sy sit', was formed from *sit*, where the *e/i* diiference originates from Ablaut; the same is true for the pair *drink/drench*: *drench* originally meant 'to make sy drink'.

4.3.5 Vocabulary: the Indo-European heritage in English

Although English has borrowed a huge amount of words from other languages, much of its basic vocabulary consists of words inherited from the ancient PIE parent language. Without providing a long list, I give you some examples, grouping them according to meaning and grammatical category. Ancient words of PIE origin include:

1. Most pronouns as well as articles, e.g. *I, you, we, this, that, my, your, who, what, the, a(n)*, etc.
2. Several prepositions, e.g. *in, at, of, to, on, over,* etc.
3. Several auxiliaries, e.g. *will, can, may* (also *would, could, might*, which are originally the past tense forms of *will, can, may*), etc.
4. Numerals from 1 to 10, as well as *hundred* [25] and *thousand*; other numerals (such as *eleven, twelve, thirteen, twenty*, etc., were coined later, though from PIE morphemes!).
5. Nouns denoting basic notions, such as body parts (e.g. *head, foot, tongue, eye, ear*), natural phenomena (e.g. *water, fire, wind, snow, moon, sun, star*), certain family relationships (e.g. *mother, father, brother, son, daughter*), everyday tools and objects (e.g. *wheel, yoke, stool, door, thatch*)[26], and several other common nouns (e.g. *feather, milk, wolf, ford, night, man*).

[25] More precisely, the *hund* part – the element *–red* is a later addition, but in Old English, *hund* is still used.

[26] The word *yoke* (Hungarian *iga*) may not seem to be a very basic word nowadays, but in earlier times, it was a common tool. *Thatch* originally meant 'roof'.

6. Verbs denoting basic activities or states, e.g. *be, come, sing, bear, think, know, work, love, see,* and many others.
7. Several common adjectives, e.g. *young, thin, full, deep*, etc.

Of course, as I said, there are many other words – the above is just an illustration. Needless to say, these words have related words, that is, cognates, in several different branches of Indo-European. The word *brother*, for example has cognate forms all around the place, cf. Latin *frater*, Czech *bratr*, Welsh *brawd* (without a final *r*, but it is still present in the plural form *brodyr*), etc. These cognate words still sound quite similar, but this isn't always the case. The English word *wheel* (Old English *hweohol*) is regularly related to Sanskrit *čakras* as well as Greek *kyklos* or Slovene *kolo* (all meaning 'wheel'), and a common PIE ancestral form can be reconstructed (something like *$*k^wek^wlos$*), from which the given related words can be unproblematically derived by regular sound changes – although at first sight, you probably wouldn't even think of these words as being related! Relatedness, however, does not mean similarity: it means common ancestry. This is the same as in the case of related people: you may look similar to a relative of yours (say, your cousin), but that's not always the case, and the more distant the relationship is, the less likely you are to look similar. Sanskrit, Latin, Slovene, Greek and English are relatives, but quite distant ones – their common ancestor, PIE, was spoken thousands of years ago. As for the word *wheel* and its cognates, the PIE ancestral form has undergone quite different sound changes in the daughter languages that it needs proper historical investigation to establish their relatedness. This point will be further illustrated by the most famous sound change in the history of Germanic languages – it is this branch of Indo-European, including English and its closest relatives, to which we turn our attention in the next chapter.

Suggested reading

Indo-European is undoubtedly the best studied language family in the world; accordingly, the literature on it, as well as on PIE, is huge. First, check Part I of the Bibliography: general introductions to the history of English all contain material on the topic. Textbooks on Indo-European languages and PIE include Clackson (2007), Baldi (1983), and Fortson (2011); Szemerényi (1999) originally published in German in 1970, is still a classic, though it does not make an easy reading.

5 Germanic: Old English and its closest relatives

5.1 Introduction

This chapter discusses the Germanic branch of Indo-European as well as some of the characteristic features of Proto-Germanic and the ancient Germanic languages in general. Unusually for a textbook on the history of English, I will also discuss Old English here, rather than devoting a separate chapter to it. The primary reason for this choice has been that Old English is still very close to the other "old" Germanic languages, and – both in its vocabulary and grammar – it is much closer to Proto-Germanic than Modern English. In other words, Old English can be used (at least in most ways) as an illustration of the basic properties of Proto-Germanic, too. There is a secondary reason, too, however: most students of English, studying from this book, are unlikely to read Old English texts in the original, so a detailed discussion of Old English seems to be beyond the scope of this book.

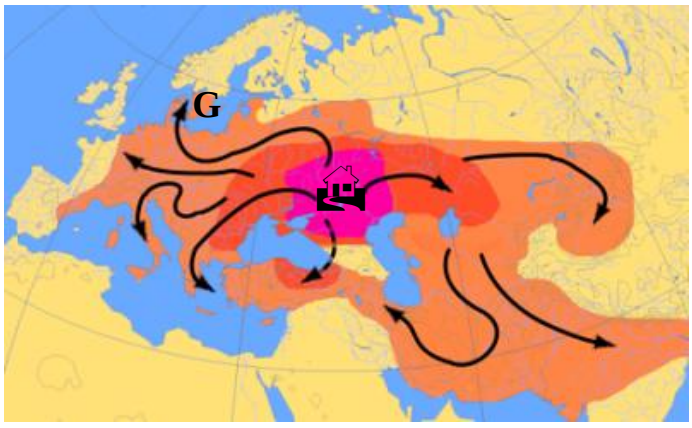
The present chapter is structured as follows. First, we present the Germanic family and its members. Second, a brief description of the chief properties of Proto-Germanic, the parent language of the Germanic family, is given. Third, Old English is discussed; the discussion will focus partly on the common Germanic properties that are still characteristic of Old English, partly on those features which have survived to some extent up to the present.

5.2 The Germanic languages: their origin and the members of the family

5.2.1 The Germanic homeland and the early spread of Germanic tribes

The Germanic languages – as discussed earlier – constitute a branch of the Indo-European family. Recall that PIE was spoken somewhere North of the Black Sea. Since the population increased, the original Indo-European homeland gradually became to small for the larger populace, and several tribes migrated to other regions. As a result of this migration, these communities became separated from each other, which resulted in linguistic divergence. In other words, the parent language became more and more different in the different areas. The picture in (64) shows the main directions of migration[27]:

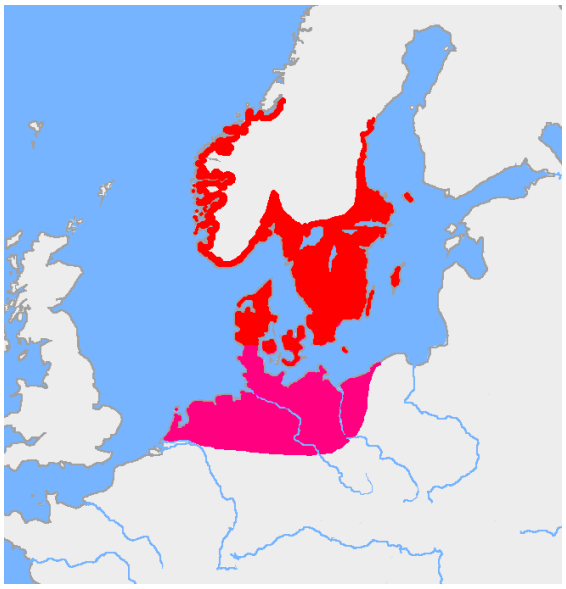
(64)



[27] Source: http://en.wikipedia.org/wiki/Proto-Indo-Europeans, accessed 22/10/2012.

The area from which the arrows start, indicated by the symbol , shows the Indo-European homeland. The arrow marked by "G" indicates the direction of the migration of the ancestors of Germanic peoples. As you can see, they moved from Eastern Europe to Southern Scandinavia, and indeed, this is the area where Proto-Germanic – formerly but a dialect of PIE – emerged gradually as an independent language. The map in (65) shows the approximate area where Proto-Germanic (from now on: PGmc) was spoken at about 500 B.C.[28]

(65)

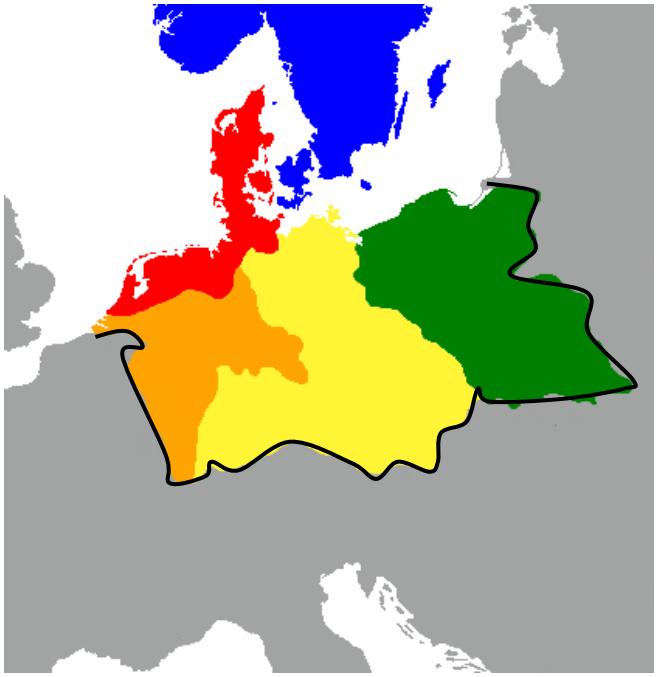


When the ancestors of Germanic peoples arrived in the area shown in (65) remains an unsettled question; it may have happened as early as 3,000 B.C., but maybe somewhat later. What seems certain is that by the first millenium B.C., this part of Europe was inhabited by Germanic-speaking tribes. The map above, as mentioned, shows the situation at around 500 B.C. At this time, Proto-Germanic was still one language: although minor dialectal differences must have existed, all Germanic speakers could happily understand each other, and the majority of changes that took place in the language affected the entire linguistic area where PGmc was spoken, so these changes did not result in diversification. Indeed, many features that characterize Germanic (as opposed to other branches of Indo-European) took place during the first millenium B.C.

Quite soon, however, several Germanic tribes started to migrate away from the Germanic homeland in all directions. Look at the map in (66)[29], showing the areas where Germanic-speaking tribes lived in the first century A.D. As you can see, the Germanic-speaking area extended to larger parts of the Scandinavian peninsula, while on the mainland of the European continent, it occupied a large area of Central Europe, bordered by the thick black line on the map.

[28] Source: http://en.wikipedia.org/wiki/Proto-Germanic_language, accessed 21/10/2012.
[29] Source: http://en.wikipedia.org/wiki/Germanic_peoples, accessed 21/10/2012, slightly modified by the author of this book; the modification is not significant as far as the content is concerned.

(66)



At this time, PGmc already showed dialectal differences, but these were still not significant enough to enable us to talk about different Germanic languages. It takes a further two to three hundred years until significant differences between dialects of Germanic can be pinpointed, but we can still state with certainty that all speakers of Germanic could understand each other's speech – at least to some degree – before about 500 A.D. (and some of them even later; indeed, Swedish and Norwegian speakers, for example, can understand each other quite well up to the present). The point is that the breakup of Germanic linguistic unity appears to have started during the first centuries A.D., but it was a gradual process.

5.2.2 The branches of Germanic and their members

After the beginning of the widespread geographical expansion of Germanic, the language became more and more diversified, and three major dialects came into existence: (i) East Germanic, (ii) North Germanic, and (iii) West Germanic.

1. The Germanic speakers whose dialect is referred to as **East Germanic** were those who – as the name suggests – migrated from the Germanic homeland towards the East. They separated from the rest of the Germanic peoples quite early, and kept migrating further on: from Central-Eastern Europe they moved on to the South-East of the continent, but some of them moved further, as far as Southern Europe (Italy) and even to South-Western Europe (the South of Gaul, i.e. present-day Southern France, as well as the Iberian peninsula. In these areas, they established kingdoms after the fall of the West Roman Empire, although these kingdoms were quite short-lived, and the East Germanic speakers themselves were quickly assimilated to the local, Romance-speaking, population. The chief representative of the East Germanic group is **Gothic**, primarily known from the surviving parts of a translation of the Bible by the Gothic bishop Wulfila from the 4th century A.D. Wulfila[30], who lived and

[30] The name *Wulfila* originates from the Gothic word *wulfs*, meaning 'wolf'; *-ila* is a dimunitive suffix (= *kicsinyítőképző*), so *wulfila* means 'little wolf'.

worked in the Balkans, created a special alphabet that he used to write Gothic, based mostly on the Greek alphabet, but he also adopted some letters from the Roman script. Most of the surviving Gothic texts are passages from the New Testament. You find a sample in (67) below:

(67)

Original:	𐌰𐍄𐍄𐌰	𐌿𐌽𐍃𐌰𐍂	𐍆𐌿	𐌹𐌽	𐌷𐌹𐌼𐌹𐌽𐌰𐌼	𐍅𐌴𐌹𐌷𐌽𐌰𐌹	𐌽𐌰𐌼𐍉	𐍆𐌴𐌹𐌽
In Roman letters[31]:	*Atta*	*unsar*	*þu*	*in*	*himinam*	*weihnai*	*namo*	*þein*
English (literally):	Father	our	you	in	heavens	be hallowed	name	yours

I.e., 'Our father, who are in heaven, may your name be holy', to give a present-day English translation. As you can see, Wulfila's Gothic alphabet is not the same thing as the version of the Roman script used in Germany for a long time, also known as Gothic, but which has nothing to do with the Goths.

Gothic – as well as other, minor, East Germanic languages – became extinct by the Middle Ages, although it continued to be spoken in the Crimean peninsula up to about the 17th century. In other words, East Germanic is by now totally extinct, so no present-day Germanic language belongs to this branch.

2. The early period in the history of **North Germanic** languages is referred to as **Old Norse**, spoken by those Germanic tribes who remained in Scandinavia rather than migrating to other parts of the continent. They are traditionally referred to as **Vikings**, who – being sea-faring people – sailed to Iceland and the Faroe Islands[32] (the language of these parts is still North Germanic), but they also travelled as far as Greenland and even to North America. Indeed, the Vikings invaded many other parts of Europe, too, including parts of the British Isles, the area in France known as Normandy, but they also went eastward, where they played a key role in establishing the Russian state, and to Southern Europe, as far as the South of Italy and Sicily. Apart from Iceland and the Faroe Islands, however, the language of the invading Vikings has not survived: the invaders were soon assimilated to the local population, linguistically speaking. (Although in some places, most notably in England, they left their impact on the local language: Old English borrowed a large amount of words from North Germanic. More details on this below.)

Old Norse remained quite a unified language for a long time: up to the first centuries of the 2nd millenium A.D., it showed but minor dialectal differences. After that, however, its diversification took more speed, and today, several North Germanic languages can be distinguished, all spoken in Scandinavia and in the North Atlantic area. On the continent, this branch is represented by **Danish**, **Swedish**, and **Norwegian**, still exhibiting a remarkable similarity between themselves. In the Northern Atlantic, **Icelandic** is spoken in Iceland, while **Faroese** in the Faroe Islands. These two languages are quite similar to each other, but quite different from the rest of North Germanic. **Danish** is spoken chiefly in Denmark, but it is also an official language in the Faroe Islands and Greenland, both of which are autonomous

[31] The letter *þ* is not really a Roman letter, but it is used in old Germanic writing, such as in Old English. As a result, it is also used when rendering Gothic texts in the Roman alphabet. It stands for the voiceless dental fricative **θ**, as in English *<u>th</u>ink*.

[32] A group of small islands in the North Atlantic, situated north of Scotland, roughly halfway between Norway and Iceland. It is an autonomous region of the Kingdom of Denmark.

regions under the rule of the Danish Crown. **Swedish** is the national language of Sweden, but it is also spoken in parts of Finland, where it is a nationwide official language alongside Finnish (even though Swedish speakers represent a small minority in Finland!). **Norwegian** is the national language of Norway. Since Norway was under Danish rule for centuries, the language of the upper classes was heavily influenced by Danish (which is close to Norwegian anyway, and was used as the official written language). During the 19th century, a new standard form known as *Nynorsk* (= 'New Norwegian') was created, but the older one, close to Danish and called *Bokmål* (= 'book language'), is still in widespread use; as a result, there are two, officially recognized written variants of Norwegian today, while Norwegians tend to use their own dialect in speech.

Faroese is the official language of the Faroe Islands, alongside Danish. As mentioned, it is similar to **Icelandic**, the national language of Iceland. Both languages, especially Icelandic, are distinguished by their remarkably conservative (archaic) features. A modern Icelandic speaker, for example, can easily read Old Norse texts! Indeed, the most significant part of Old Norse literature was written in Iceland, including the poetic and prosaic texts collectively known by the term *Edda*, as well as the prosaic texts known as *sagas* (in fact, the word *saga* – originally meaning 'something told/said' – found its way into many European languages, including English and Hungarian). These are the best known texts in Old Norse, but there are many others – indeed, Old Norse literature is enormous.

3. The **West Germanic** branch of the Germanic family consists of several languages which can be classified according to a number of criteria. Since any classification within West Germanic involves lots of argumentation and is also problematic, I leave the problem aside; students who are interested in the topic are referred to the Suggested Reading section at the end of this chapter. Nonetheless, cases of close relatedness will be pointed out.

The West Germanic languages were originally spoken in what is present-day Denmark, the Netherlands, and Northern Germany. Gradually, they spread in two main directions: (i) towards the South-Central areas of Europe, as far as the North of Italy, (ii) to Britain. Let us first see those West Germanic languages which are (as indigenous languages) spoken on the Continent – in other words, all of them excepting English.

As far as the number of speakers is concerned, the largest continental West Germanic language is **German**. It is the national language of Germany and Austria, and one of the national languages of Switzerland. It is also spoken in several neighbouring countries (such as Italy, Belgium, Luxemburg or Hungary), not to mention German emigrant communities in many countries, chiefly in the US. The estimated total number of native German speakers is close to 100 million, making German not only the largest continental Germanic language, but also the second largest (after English) all over the world.

That said, it must be emphasized that when we talk about German, we actually mean two rather different things. First, the area where German is spoken is highly diverse dialectally speaking. The Swiss dialects, to mention an extreme example, are practically unintelligible to most other speakers, but it is generally true that southern dialects are quite different from northern ones. This is because the northern dialects of German, known as **Low German** because they are spoken in the lowlands close to the North Sea and the Baltic Sea, originate from the old West Germanic language called **Old Saxon**, which was actually more closely related to English than southern German dialects! The dialects spoken in the middle and southern areas are called **High German** (because they are spoken in the highlands, i.e. the mountainous parts). Linguistically, old High German dialects were rather different from Low German ones. What we mean by "German", then, is historically a bunch of loosely connected West Germanic dialects. The reason why they are all called German is that due to cultural and political reasons, they use a common standard: in fact, this is the other sense of the word "German" – **Standard German**, which is based on High German dialects spoken in

the middle of Germany. Due to its political and cultural prestige, it has by now replaced most Low German dialects, which have been regarded as rural and uneducated. Indeed, the number of Low German speakers is rapidly decreasing: while no fully reliable statistics are available, the number of native Low German speakers in Germany is estimated to be around 3 million, who are mostly elderly people living in rural areas.

Closely related to Standard German is **Yiddish**, the traditional language of Ashkenazi Jews in Central and Eastern Europe before World War II. Yiddish, though written in the Hebrew script, is a West Germanic language, more precisely, the descendant of medieval South-Western High German dialects. Specifically, Jews living in the area known as the Rhineland in the early part of the Middle Ages spoke the same language as everyone else who lived there: a dialect of High German. Later on, however, migration towards Eastern Europe began, and the language was carried along, too. Being separated from other High German areas, it started its independent development, gradually becoming distinct from German, but still retaining a high degree of similarity. Before the 1940's, Yiddish was widely spoken over vast areas covering practically all of Central and Eastern Europe, but also in North America, by Jewish immigrants. On the eve of World War II, the total number of speakers was over 10 million. Following the Holocaust, however, this number tragically fell; massive assimilation also contributed to the decrease, as well as the fact that Yiddish has been strongly disfavoured in the State of Israel, where the Zionist movement has always propagated the use of Hebrew. As a result, the total number of Yiddish speakers may now be under 2 million, most of whom live in the US.

Closely related to Low German is **Dutch**, spoken in the Netherlands and the Northern part of Belgium, known as Flandria, where it is called **Flemish** (but the two terms refer to the same language; I use Dutch here for the sake of simplicity). Dutch originates from an old West Germanic dialect known as **Old Low Franconian**, called so because it was originally spoken by the part of the Franks who settled in the lowlands along the river Rhine[33]. Dutch has also been carried overseas by colonization, so it is spoken in several parts of the world, though not by a significant amount of people. Dutch emigration to Southern Africa, however, gave rise to the birth of a new West Germanic language, called **Afrikaans** (which simply means 'African'). Afrikaans is spoken by a significant proportion of the population of the Republic of South Africa, where it is one of the many official languages, but it is also spoken and widely understood in neighbouring countries. Originally, Afrikaans was a variant of 17th century Dutch, carried to South Africa by Dutch settlers, but it started to develop differently, and today it is considered to be an independent language – although still highly similar to Dutch.

Last but not least, there is another West Germanic language to be mentioned: **Frisian**. It is chiefly spoken in the North of the Netherlands, but also along the North Sea coast in Germany. The total number of speakers is below half a million, most of whom live in the Netherlands, but it must be emphasized that Frisian is historically not very close to Dutch. Indeed, its closest relative is – take heart – English! This is because most of the Anglo-Saxon invaders of Britain came from the North Sea coast, where the ancestors of modern Frisians also lived; at the time when the Anglo-Saxons left for Britain, their language was practically identical to the language of the Frisians (in other words, Frisians may well be considered to be relatives who were "left behind" on the continent). Indeed, Old Frisian is so similar to Old English that a knowledge of either of them will enable you to read a text in the other without any serious difficulty.

[33] The other part of Franks settled in the Southern areas along the Rhine, and they spoke a High German dialect. It was them who played the central role in establishing the Frankish Empire, and also the ones who gave their name to the country known as France, originally *Franc+ia*, i.e. the land of Franks. It is another story that they soon became assimilated to the Romance-speaking majority of the land.

This concludes our discussion of the Germanic family. English is still missing, but we will come back to it soon; first, however, let us take a look at the chief features of Proto-Germanic, the parent language of all Germanic tongues.

5.2.3 Proto-Germanic

5.2.3.1 Phonology

PGmc is characterized by several phonological changes which set it apart from other branches of Indo-European, i.e. changes which do not take place elsewhere in the Indo-European family. On the other hand, since these changes affected PGmc, their results are present in all the Germanic languages – although later changes may have obscured the original ones.

Some of the PIE vowels undergo certain changes in PGmc, but what is truly interesting is the overall restructuring of the obstruent system. Recall the system of obstruents reconstructed for PIE, repeated here in (68):

(68) PIE obstruents

Place of articulation →	Labial	Dental	Alveolar	Velar	Labio-velar
Fricatives			**s**		
Voiceless stops	**p**	**t**		**k**	**k^w**
Voiced stops	**b**	**d**		**g**	**g^w**
Voiced aspirated stops	**b^h**	**d^h**		**g^h**	**g^{hw}**

Let me remind you of two interesting features of this system: (i) there is only one fricative, the voiceless alveolar **s**, (ii) but, by contrast, there is a rich system of stops (12 altogether).

This system is inherited by PGmc, which then changes it substantially: all the stops undergo some change, and not in a random way, but following a nice pattern – moving around a circle, as it were. The change is known as *Grimm's Law*, named after Jacob Grimm (of fairy-tale fame – but he was also a linguist), who was the first to effectively publicize it in the early 19th century (it was not discovered by him, though, but by a Danish linguist, Rasmus Rask, some years earlier).

1. Step 1 of Grimm's Law: voiceless stops become voiceless fricatives, that is:

p > **f**
t > **θ**
k > **x**
k^w > **x^w**

Note: **x** represents a voiceless velar fricative, as in Hungarian *doh,* rather than **ks**!

After this move, the system in (68) looks as follows:

(69) The PGmc obstruent system after Step 1 of Grimm's Law

Place of articulation →	Labial	Dental	Alveolar	Velar	Labio-velar
Fricatives	**f**	**θ**	**s**	**x**	**x^w**
Voiceless stops					
Voiced stops	**b**	**d**		**g**	**g^w**
Voiced aspirated stops	**b^h**	**d^h**		**g^h**	**g^{hw}**

There are now four new fricatives – indeed, all boxes in the fricative row are filled, but voiceless stops are gone, as shown by the shaded boxes.

2. Step 2 of Grimm's Law: voiced stops become voiceless stops, that is:

b	>	**p**
d	>	**t**
g	>	**k**
g^{w}	>	**k^{w}**

This move results in a further shift: the shaded boxes in (69) are re-filled, but this time, the voiced stop row is left empty:

(70) The PGmc obstruent system after Step 2 of Grimm's Law

Place of articulation →	Labial	Dental	Alveolar	Velar	Labio-velar
Fricatives	**f**	**θ**	**s**	**x**	**x^{w}**
Voiceless stops	**p**	**t**		**k**	**k^{w}**
Voiced stops					
Voiced aspirated stops	**b^{h}**	**d^{h}**		**g^{h}**	**g^{hw}**

Now comes the "final cut":

3. Step 3 of Grimm's Law: voiced aspirated stops become unaspirated, that is:

b^{h}	>	**b**
d^{h}	>	**d**
g^{h}	>	**g**
g^{hw}	>	**g^{w}**

This final move restores the voiced stops, so the shaded boxes in (70) are re-filled, while the aspirated stops disappear: there is no row below to re-supply them! The result is:

(71) The PGmc obstruent system after Step 3 of Grimm's Law, i.e. after Grimm's Law is completed

Place of articulation →	Labial	Dental	Alveolar	Velar	Labio-velar
Fricatives	**f**	**θ**	**s**	**x**	**x^{w}**
Voiceless stops	**p**	**t**		**k**	**k^{w}**
Voiced stops	**b**	**d**		**g**	**g^{w}**

Although voiced aspirated stops are gone for good, please count the total number of obstruents in (71) – and then do the same for (68). You will find that the number of obstruents hasn't changed: four of them have disappeared, but four new ones have been born! What really happened is that the voiceless stops started to move away from their original place, giving rise to a sort of "chain reaction" – no wonder that such shifts are called *chain shifts* in historical linguistics!

I must confess, though, that the overall picture is not as neat as suggested here. Aspirated stops, for example, turn to voiced fricatives in some positions, but I do not believe that a very detailed discussion is really necessary: if you are interested in the minor details,

please check the Suggested Reading section. What I find important, however, is to note that the velar fricative **x** – as well as its labiovelar peer, **x^{w}** – turn to **h** (**h^{w}**) in many cases, most notably, at the beginning of words.

Let us now see some examples, however – after all, so far this whole Grimm's Law business has been but a fine intellectual exercise, but what's the use? Here we go, then! Each example will start with a PIE form, then the English form is given, illustrating the operation of Grimm's Law (since English is Germanic). Finally, one or two non-Germanic cognates are given: the aim of these is to show that Grimm's Law does not take place outside Germanic, so the given forms preserve the original PIE stop.

(72) Grimm's Law: examples (note: E = English, OE = Old English)

(a) PIE voiceless stops > PGmc voiceless fricatives

PIE > Gmc	**Compare:**
PIE *pətər 'father' > E *father*	Latin *pater*, Sanskrit *pitar*
PIE *trejes '3' > E *three*	Latin *tres*, Welsh *tri*
PIE *kerd 'heart' > E *heart*	Greek *kardia*, Latin *cordem*
PIE *k^{w}od 'what' > OE *hwæt* (E *what*)	Latin *quod*

Notes: (a) Latin *c* = [k], *qu* = [k^{w}], (b) OE *hwæt* = [hwæt], note that the modern spelling has *wh*, instead of OE *hw*, but this is simply a spelling convention: recall from the discussion in Chapter 3 that in EModE, word-initial *wh* was still pronounced [hw] – as it is still in several English accents, such as Scottish English.

(b) PIE voiced stops > PGmc voiceless stops

PIE > Gmc	**Compare:**
PIE *abl 'apple' > E *apple*	Russian *jabloko*, Welsh *afal*
PIE *dwo: '2' > E *two*	Latin *duo*, Russian *dva*
PIE *genos 'kin, kind' > E *kin*	Latin *genus*, Greek *genos*
PIE *g^{w}en 'woman' > E *queen*	Greek *gyne* 'woman'

Notes: (a) the *–ko* in the Russian word is originally a suffix, (b) Welsh *f* = [v], a regular development of Proto-Celtic [b], (c) the Greek word *gyne* is pronounced [gynɛ:], from earlier [g^{w}ynɛ:].

(c) PIE aspirated stops > PGmc voiced stops

PIE > Gmc	**Compare:**
PIE *b^{h}er 'to bear' > E *bear*	Sanskrit ***bhar**-ami*, Latin ***fer**-o* 'I bear'
PIE *d^{h}e:/d^{h}o: 'to do' > E *do*	Sanskrit *da-**dha**-mi*, Latin ***fa**c-io* 'I do'
PIE *g^{h}est 'foreigner' > E *guest*	Latin ***host**-is* 'enemy'

Notes: (a) no example of PIE ghw is given, because its ultimate developments are so complex that it would need too many explanations to make the picture clear, (b) in the Sanskrit/Latin forms, consider the boldfaced parts (the rest are affixes), (c) the fricatives **f** and **h** are regular developments of aspirated stops in Latin, (d) note the interesting meaning changes in the case of PIE *g^{h}est 'foreigner'.

Grimm's Law is of special relevance for Modern English primarily because English has borrowed a huge amount of words from Latin and Greek. This often results in *doublets*,

when you have two words: one of Germanic origin (showing the effects of Grimm's Law), and another one, borrowed from Greek or Latin, having the same PIE root, but without the effects of Grimm's Law. Here are some examples:

(73)	*father – **pater**nal*	The latter is from Latin, which preserves the **p**
	*two – **du**al*	The latter is from Latin, which preserves the **d**
	*heart – **card**iology*	The latter is from Greek, which preserves the **k** as well as the **d**
	*guest – **host**ile*	The latter is from Latin, where $\mathbf{g^h} > \mathbf{h}$, cf (72c).

There are hundreds of further examples; if you are interested, check a good etymological dictionary.

5.2.3.2 Morphology

Many of the morphological features of Proto-Germanic were mentioned in the previous chapter. Let us now sum up the most important points.

1. As for nouns, adjectives, and pronouns, the Ablative and the Locative case is lost, their functions taken over by the Dative. The Instrumental still survives, but is well on the way to be lost, and the same is true for the Vocative. Altogether, only four cases (Nominative, Accusative, Genitive, Dative) remain stable. The Dual is still present, but again, it is becoming extinct: nouns, for example, no longer distinguish the Dual from the Plural in Gothic, the earliest extensively documented Germanic language. On the other hand, adjectives become more complex. In PIE, they were inflected in the same way as nouns, but in PGmc, another pattern came into existence: as a result, adjectives could be inflected in two different ways, depending on their syntactic function. I will use the (untraditional) terms **indefinite** and **definite** to refer to these two ways of inflection[34]. The indefinite inflection was used when the adjective in attributive function was preceded by a determiner (such as a definite article, a demonstrative or a possessive pronoun), while the definite inflection was used in other cases, i.e. when it had a predicative function, or when its function was attributive, but there was no determiner before it. Illustrations are given in (74) below, taking examples from Old English:

(74)	Indefinite:	(a)	*þā gōdan menn* 'the good men'
	Definite:	(b)	*gōde menn* 'good men'
		(c)	*þā menn sind gōde* 'the men are good'

In (74a), the indefinite form, taking the suffix *–an*, is used: the attributive adjective is preceded by the definite article. In (74b), there is no determiner, so the definite form, taking the suffix *–e* is found, and you have the same form in (74c), where the adjective is part of the predicate. You find a similar (though not exactly identical) pattern in modern German, cf. *guter Mann* 'good man' (definite) vs. *der gute Mann* 'the good man' (indefinite).

2. The system of verbs undergoes radical changes, chiefly in the direction of simplification. For one thing, the Optative mood is completely lost as a distinct mood, its

[34] They are called, traditionally, **weak** and **strong**, respectively. I avoid these terms here because they are also used for verbs (see below), in a radically different sense.

functions being "swallowed" by the Subjunctive. The original Middle Voice forms are getting lost, too. Finally, the relatively complex Aspect/Tense distinctions of PIE are also simplified: the Perfect/Imperfect/Aorist division is gone, and PGmc verbs distinguish two tenses only: Present and Past. Related to this, however, is a development which enriches the language: the emergence of a new Past Tense formation, characteristic of Germanic and not found in other branches of Indo-European. Let us now see this development, especially because it underlies a fundamental difference between the Past Tense formation of two major classes of verbs up to the present.

Remember the discussion of Ablaut in the previous chapter, where the phenomenon was illustrated – among other things – by verbs such as *sing, drive, come*, etc., whose Past Tense forms display different vowels than the Present Tense form; indeed, it is the vowel alternation only which distinguishes Present and Past, cf. *sing – sang, drive – drove, come – came*, etc. This vowel difference (Ablaut), then, goes back to ancient, PIE, times. Verbs which use Ablaut to express Past Tense are traditionally called **strong verbs** in Germanic linguistics. In PGmc, however, a new way of expressing Past Tense came into existence, whereby a suffix was added to the verb. This suffix contained the consonant **d**, sometimes **t**, and it was added to the verb without any modification in the vowel of the root. Such verbs are traditionally called **weak verbs**, and they soon became widespread, serving as the regular pattern of Past Tense formation in the Germanic languages. Indeed, all regular verbs in Modern English are historically weak, including, for example, *play – played, kiss – kissed, load – loaded*: in all cases, there is a suffix showing Past Tense, and the vowel of the root is unchanged.

It is not true, however, that all verbs which are historically weak are necessarily regular, too. This is because many weak verbs underwent sound changes during and after the PGmc period, which have resulted in a vowel difference between the Present and the Past Tense forms. Some of these sound changes are quite ancient, and occurred already in PGmc, resulting in highly different forms; such verbs include *bring – brought, think – thought*, etc. If you compare the German cognates of these forms, you can discover a similar pattern, cf. German *bring – bracht, denk – dacht*, etc.[35], showing that the irregular pattern goes back to PGmc times, since it is found in other Germanic languages, too, not only in English. At the same time, you can still detect a Past Tense suffix *t*, absent in the Present Tense forms; compare strong verbs, where there is no such suffix (cf. *sang, drove, came*). Many other changes, however, took place in the individual history of English, during the Old and Middle English periods, resulting in irregular weak verbs such as *keep – kept, dream – dreamt, mean – meant, sell – sold, tell – told*, where there is a vowel difference again, but the Past Tense suffix *d/t* (underlined) is still recoverable. In other words, in Modern English,

(a) All strong verbs are irregular, and
(b) All regular verbs are weak, but
(c) Not all weak verbs are regular.

There is a small group of weak verbs which used to contain a Past Tense suffix, but it disappeared, cf. *meet – met, feed – fed*, etc. These look like strong verbs, since you cannot identify a Past suffix. Yet, in earlier English, there *was* a suffix. Such verbs will be discussed in the next chapter.

[35] The German forms are generally followed by a suffix indicating Person/Number, but I omit these for the sake of simplicity, and I only give the roots.

5.2.3.3 Vocabulary

While the major part of the vocabulary of PGmc is inherited from PIE, an estimated 30% of its word stock is of uncertain origin. These items are found in all (or at least most) Germanic languages, but they have no cognates in other branches of Indo-European. The most probable reason is that PGmc speakers came into contact with some non-Indo-European language when they settled in Southern Scandinavia thousands of years ago; they borrowed many words from this unknown language, which later became extinct (but this explanation, while probable, is still a matter of conjecture: there is no decisive proof either for or against it). In Modern English, there are several words, too, which belong here, including (without providing a long list) *hand, sea, drink, bath, wife, rain, ship*, etc. (cf. German *Hand, See, trinken, Bad, Weib, Regen, Schiff*).

PGmc speakers, of course, also came into contact with Latin, from which they borrowed several words, mostly items which are related to Roman or Mediterranean culture, including *wine, street, kitchen, dish*, etc. Other languages, such as Greek or Celtic, also contributed to the enrichment of the PGmc word stock, though to a lesser extent than Latin; the most notable Greek borrowing into PGmc is probably *church*, originally sounding something like **kirike*, from Greek *kyriakon*; in Old English, the original **k**'s turned to **tʃ**, hence the modern form, but in the far North, they remained – hence the Scottish word *kirk*, used today to refer to the Church of Scotland! (Compare also German *Kirche* 'church'.)

5.3 Old English

5.3.1 The external history of Old English: a brief sketch

According to historical tradition, originating from Bede's (6th century) *The ecclesiastical history of the English nation*, three West Germanic tribes, the **Angles**, the **Saxons**, and the **Jutes** arrived in Britain in 449 A.D., from their original homelands in the North-West of present-day Germany and Southern Scandinavia. Tradition holds that they were invited by the then king of the Britons (Celtic inhabitants of former Roman Britain), who needed assistance in their fight against the Picts, who kept attacking them from the North, i.e. present-day Scotland. The reason for this invitation was that during the early years of the 5th century, the last Roman troops left Britain, and the Britons were left alone. Tradition holds that after the three West Germanic tribes arrived in Britain, they defeated the Picts all right, but then they turned against the Britons, and eventually conquered most of present-day England (with the exception of Cornwall) and the South of Scotland. As a result, Britons were forced to flee to Wales, Cornwall, and some of them migrated to Brittany. The Germanic invaders are collectively called **Anglo-Saxons**, since the Jutes were small in number compared to the Angles and the Saxons. Most of England was settled by the Angles, who gave their name to the language and the country.

How much of this tradition is true is a matter of debate – for example, it is certain that the Anglo-Saxon invasion cannot be precisely connected to a single year; indeed, some Germanic immigrants settled already in Roman Britain in the 4th century, and the massive settling of the Anglo-Saxons took several decades after the withdrawal of the last Roman troops from Britain. At any rate, we know almost nothing of the period between about 450 and 600. What seems certain is that by the 7th century, the Anglo-Saxons had established several kingdoms in Britain, including Northumbria, Mercia, Kent and Wessex. During the later part of the 7th century as well as most part of the 8th, Northumbria (corresponding to

modern Northern England and Southern Scotland, called Northumbria because it was the area north of the river Humber) was the most powerful of the Anglo-Saxon kingdoms, with Mercia (corresponding to the area called the Midlands today) emerging as a strong rival during the middle part of the 8th century. But then, during the last years of the 8th century, the dominance of Northumbria and Mercia came to a sudden end: they were invaded by Vikings (whom the Anglo-Saxons called Danes, though most of them might have been Norwegians), and for about a century, most of England came under Viking rule. The centre of English power shifted to Wessex, which had not been conquered by the Vikings (the term *Wessex* is used here to refer to the area lying south of the River Thames excepting Kent). During the reign of King Alfred the Great in the late 9th century, Wessex successfully resisted further attempts at invasion by the Vikings, and it produced a wealth of literature written in (Old) English. Winchester, the capital of Wessex, became also the main English cultural centre. By the 10th century, most the Viking invaders had permanently settled in Northumbria and Mercia, gradually mixing with the English population; those who had not settled had left the land. As a result, under the leadership of Wessex, a unified Anglo-Saxon kingdom was established during the 10th century, which existed up to 1066, when the Norman Conquest took place.

As mentioned above, Old English literature started to flourish in Wessex in the 9th century. The consequence of this is twofold. First, the dialect of Wessex, called West Saxon, came to be used as a pan-English standard, which was the most general form of written English up to the Norman Conquest. Second, the overwhelming majority of Old English texts is in the West Saxon dialect, and it is this variant of Old English which is usually presented in Old English grammars and textbooks. It must be emphasized, though, that several texts written in other dialects survive as well.

The major early kingdoms – Northumbria, Mercia, Wessex and Kent – also coincide with the major dialects of Old English, shown on the map below:

(75) The dialects of Old English[36]

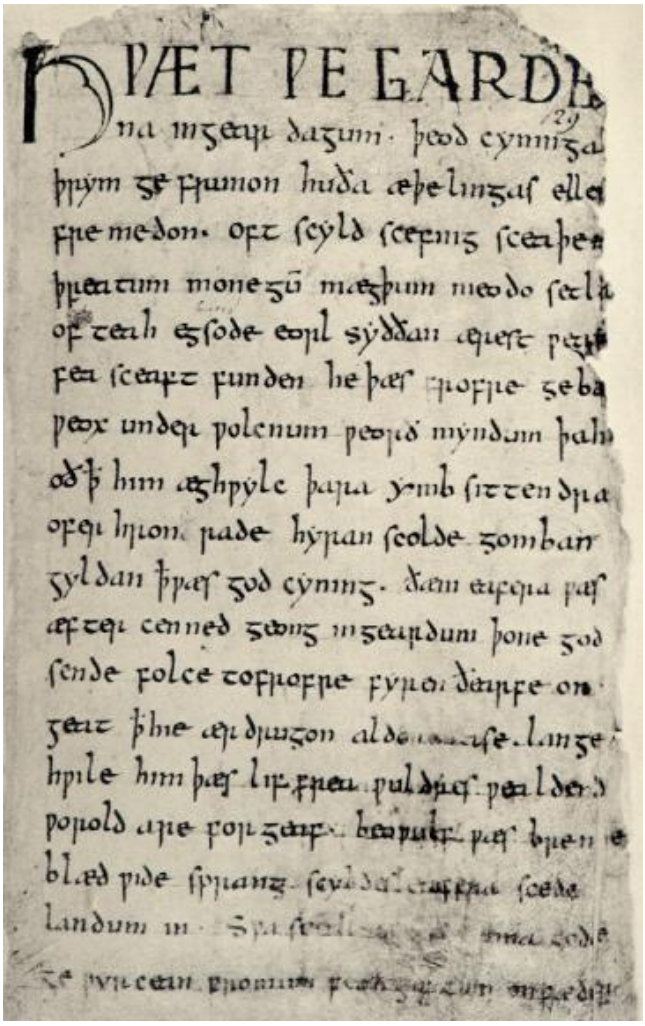
[36] Source: http://www.yorkshiredialect.com/Oegen.htm, accessed 22/10/2012.

According to tradition, by the way, Wessex was originally populated by the Saxons, Kent by the Jutes, while Northumbria and Mercia were settled by the Angles. Indeed, Mercian and Northumbrian show a lot of common dialectal features, which is why the two are collectively called **Anglian** dialects. The dominance of Angles – both in terms of population and of geographical extension – resulted in the fact that the language of the entire country came to be known as *English*, from the word *Engle*, meaning 'Angles' in Old English; indeed, by the time West Saxon became the standard, the term *English* had become so much established that it was commonly used for the language (as well as the nation) even in Wessex, and has remained in use ever since. The Angles also gave the name to the country: the name *England* comes from the Old English phrase *Engla land*, where *Engla* is the Genitive form of *Engle*, so it meant 'land of Angles', shortened over the centuries to *England*.

Though English literacy started to flourish in the 9th century, our earliest documents written in English date from the 8th century. The literature written in Old English is enormous – not only in quantity, but also regarding the variety of genres: from the great epic poem *Beowulf* (probably the best known masterpiece of Old English literature) through elegies and other poems and funny riddles to prosaic works such as the *Anglo-Saxon Chronicle*, a year-by-year chronicle describing the most important events in early English history, or several magnificent sermons. Besides, a great amount of translations of foreign texts also survives. It would be impossible to list even the most outstanding pieces of literature here. As a point of interest, you can see the first page of the Beowulf manuscript in (76)[37]:

[37] Source: http://en.wikipedia.org/wiki/Old_English, accessed 22/10/2012.

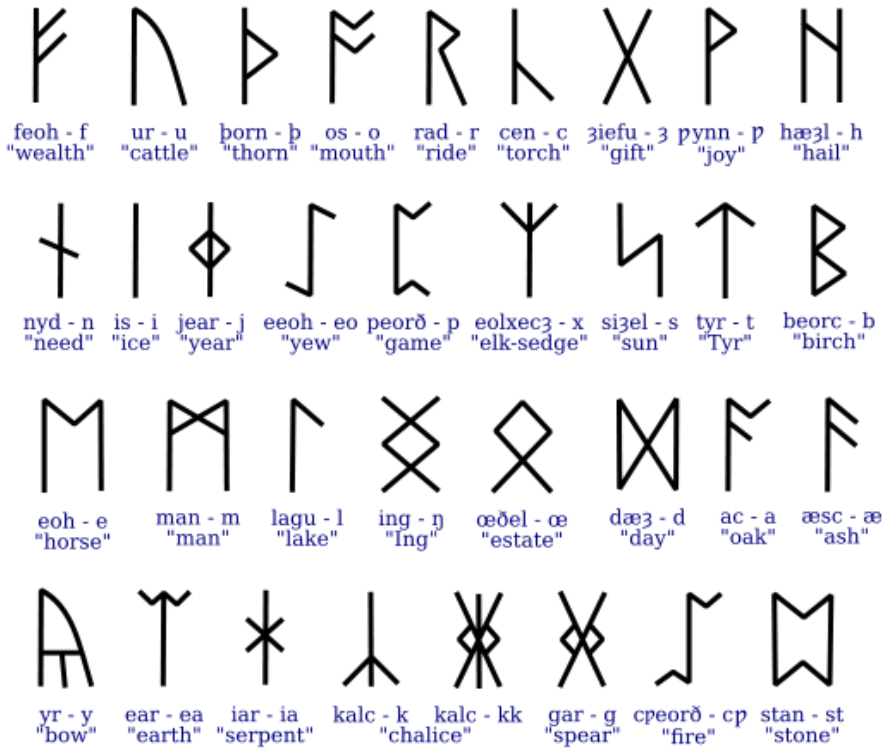
(76)



As you can see, the writing is basically Roman, although with some letter forms which seem strange to us, modern readers (which, however, was not unusual in the Middle Ages). There was also another alphabet used in Old English times, originating from an old form of Germanic script: the so-called **Runic** alphabet, also known as the **Futhark**, similar in form as well as function to the old Hungarian *rovásírás*. Much like *rovásírás*, the Runic alphabet was used to write short inscriptions, carved in stone or wood. Contrary to popular misbelief, the Runic alphabet was *never* used to write longer texts (neither was *rovásírás*): all around the Germanic-speaking world, the Roman alphabet has always been used for such purposes. Nonetheless, some Runic letters were adopted and became members of the Old English Roman alphabet; the most notable one is the letter *þ*, called *thorn*, used in Old English to denote the dental fricatives **θ** and **ð**, sounds which did not exist in Latin. (The modern spelling *th* for these sounds originates from the Middle English period.) Thus, *thorn* was written *þorn*, for example, in Old English. As a point of interest, though, I give a picture below, showing the letters of the Futhark, i.e. the Runic alphabet[38]:

[38] Source: http://en.wikipedia.org/wiki/Old_English, accessed 22/10/2012.

(77)



This point leads us to our next topic: an overview of the most important properties of Old English.

5.3.2 Old English: The language itself

5.3.2.1 A sample text with explanations

Let us now see a short text in OE, a passage from the Gospel according to St. Matthew, about the wise man who bulit his house on a rock (I hope you are familiar with it).

(78) A sample text in OE: Matthew 7:24-25

24 Ælč þāra þe þās mīn word ğehīerþ, and þā wyrcþ, biþ ğelīč þǣm wīsan were, sē ğetimbrode his hūs ofer stān.

25 Þā cōm þǣr reğen, and mičel flōd, and þǣr blēowon windas, and āhruron on þæt hūs, and hit nā ne fēoll: sōþlīče, hit wæs ofer stān ğetimbrod.

(79) A ModE translation[39]

24 Each of those who hears these words of mine, and does them, is similar to the wise man, who built his house on stone.

25 Then there came rain, and a big flood, and there blew winds, and fell down on the house, and it did not fall: truly, it was built on stone.

For the sake of easier reading, I provide a phonetic transcription of the text (80), then some notes on pronunciation follow; after that, we'll look at the words and the grammar.

(80) The text in (78) in phonetic transcription

24 æːltʃ θɑːrɑ θe θɑːs miːn word jeˈhiːərθ ɑnd θɑː wyrkθ biθ jeˈliːtʃ θæːm wiːzɑn

were seː jeˈtimbrode his huːs over stɑːn

25 θɑː koːm θæːr rejen ɑnd mitʃel floːd ɑnd θæːr bleːowon windɑs ɑnd ɑːˈhruron

on θæt huːs ɑnd hit nɑː ne feːoll soːðliːtʃe hit wæs over stɑːn jeˈtimbrod

As you can see, OE spelling and pronunciation is quite consistent: all written letters are pronounced, and a given letter is mostly pronounced the same everywhere. Let us now see some details.

The vowel letters are often found with a **macron** (= ¯) above them. This mark indicates a long vowel, so <i> = **i**, <ī> = **iː**; in fact, all vowels can be short or long. I must add that this length marking was not used in OE texts, but it is customary to show long vowels in OE grammars and dictionaries.

The letter <æ> is pronounced **æ** (as in ModE *cat*), <y> is pronounced **y** (= Hungarian *ü*), and <a> is **ɑ** (as in ModE *park*). The other vowel letters are basically pronounced as in Hungarian (or most continental languages), rather than as in ModE.

As for consonants, the following points should be noted:

1. The letter <þ> was (as mentioned earlier) pronounced as **θ** or **ð**, but not in a random way: it was pronounced as **ð** inside a word between voiced sounds, but as voiceless **θ** elsewhere. Compare the word *sōþlīče*, where it stands for **ð**, as opposed to other instances. Another letter, <ð>, was also used in the same function, but I do not use it here for the sake of simplicity; what is important is that it was pronounced as **θ** or **ð**, too, just like <þ>, and the same word could be written with either letter without any difference in pronunciation. So, for example, *þæt* also occurs in the spelt form *ðæt*, but in both cases, it is pronounced **θæt**.

2. In a similar fashion, the letters <f> and <s> could represent voiceless and voiced consonants, too: <f> stood for voiced **v** between voiced sounds, as in the word *ofer*, but for voiceless **f** elsewhere, as in *fēoll*; <s> was pronounced voiced **z** between voiced sounds (as in *wīsan*), but **s** elsewhere (as in *hūs*).

3. In original OE texts, the letter <c> could stand for either **k** or **tʃ**, but it is often difficult to predict the pronunciation. Therefore, modern editions use a diacritic mark above the letter <c> where it represents **tʃ**; I use <č> here. In all other cases, the simple <c> is to be pronounced as **k**.

[39] My translation; it is not very elegant, but I tried to keep it as close to the OE text as possible.

4. Similarly, the letter <g> could stand for either **g** or **j**, but again, it is often difficult to predict the pronunciation. Therefore, modern editions use a diacritic mark above the letter <g> where it represents **j**; I use <ğ> here. In all other cases, the simple <g> is to be pronounced as **g**.[40]

3. Stress usually fell on the first syllable of words, with some exceptions (these are invariably words with an unstressed prefix, such as *ğe*hīerþ, where the underlined part is an unstressed prefix. See below for detailed explanations; in the phonetic transcription in (80), therefore, I only indicated stress in such cases.

Let us now see the detailed comments, word by word; then, a summary will be given in the following section

	Word	**Meaning**	**Notes**
1.	ælč	*each*	> ModE *each*
2.	þāra	*of those*	The GenPl form of the word *se*, functioning both as a demonstrative (= 'that') and a definite article (= the).
3.	þe	*who, which*	A general relative pronoun, which can be used to refer to things as well as persons
4.	þās	*these*	The AccPl form of *þis* 'this'
5.	mīn	*my*	The Neuter AccPl form of *mīn* (unchanged in form), used both in the sense 'my' and 'mine' in OE
6.	word	*words*	The AccPl form of *word* (identical to the Sg, like *sheep* in ModE); *word* is a Neuter noun, and many Neuters have a suffixless plural in OE. The form *mīn* does not take a suffix either, since it agrees with *word*. Altogether, *þās mīn word* is literally 'these my words', but this structure is impossible in PdE: instead, 'these words of mine' is used.
7.	ğehīerþ	*hears*	The 3rd person Sg Present Tense form of *ğehīeran* 'to hear'. Note the suffix *–þ* (= EModE *–th*). The prefix *ğe* indicates a completed action; Hungarian *meghall* is a good equivalent, vs. *hīeran = hall*. The suffix *–an* is characteristic of OE infinitives. From *hīeran* > ModE *hear*.
8.	and	*and*	> ModE *and*
9.	þā	*those*	The AccPl of *se* (cf. row 2. above), here referring back to *þās mīn word*.
10.	wyrcþ	*does, performs*	The 3rd Person Sg of *wyrcan*, whence ModE *work*, but in OE, it is used in a wider sense, i.e. 'do, perform'.
11.	biþ	*is*	A 3rd Person Sg form of *be*; *is* also exists in OE.
12.	ğelīč	*similar*	--
13.	þæm	*to the*	The DativeSg form of *se*
14.	wīsan	*wise*	The Indefinite DatSg form of *wīs* 'wise'. The

[40] In fact, the letter <g> could also stand for other consonant sounds, but from the point of view of the above text, this is irrelevant; I refer the interested students to detailed descriptions of OE, mentioned in the Suggested Reading section.

			Definite form would be *wīsum*, but here the Indefinite form is used because the adjective is preceded by the definite article.
15.	were	*man*	The DatSg form of *wer* ‘man’, an ancient word of PIE origin; in ModE, it survives in *werewolf*, literally ‘man-wolf’. Altogether, *þǣm wīsan were* = ‘to the wise man’.
16.	sē	*who, which*	A Masculine relative pronoun, which can refer back to any Masculine noun, though generally used for male beings. Here, it refers back to *wer*.
17.	his	*his*	The Genitive form of *hē* ‘he’.
18.	hūs	*house*	Here, *hūs* is an object, so this is an Accusative form, but the same form could also function as a subject, i.e. a Nominative form.
19.	ofer	*on, upon, over*	> ModE *over*; in OE, it is used in the sense of ‘on, upon’, too.
20.	stān	*stone*	> ModE *stone*; here, it stands after a preposition, so this is an Accusative form, but the same form could also function as a subject, i.e. a Nominative form.
21.	ğetimbrode	*built*	We have the same prefix again as in row 7. above. The prefixless form is *timbrode*, the Past Sg 3rd person form of *timbrian* ‘to build’. Note the suffix *-ode* to show Past Tense (and no Ablaut), so this is a Weak Verb. The verb does not survive into ModE, but the noun from which it was formed does (= *timber*).
22.	þā	*then, thereafter*	Not to be confused with *þā* in row 9. – this is another word which happens to sound identical.
23.	cōm	*came*	The 3rd Person Sg Past Tense form of *cuman* ‘to come’, a Strong Verb (note the vowel difference due to Ablaut)
24.	þǣr	*there*	> ModE *there*
25.	reğen	*rain*	A NomSg form (it is a subject); > ModE *rain*
26.	mičel	*big*	A NomSg form again, this time a Definite one, since *mičel flōd* ‘a big flood’ has no determiner: in OE, there is no indefinite article as yet. The Indefinite form would be used, e.g., in ‘the big flood’: *se mič(e)la flōd*. The word does not survive into ModE as a common adjective, but it is preserved as a surname: *Mitchell*.
27.	flōd	*flood*	A NomSg again; > ModE *flood*, though in OE, it can also mean ‘stream, high tide’.
28.	blēowon	*blew*	The Past Plural form of *blāwan* ‘to blow’, a Strong Verb; > ModE *blow/blew*.
29.	windas	*winds*	The NomPl of *wind*, a Masculine noun (> ModE *wind*).
30.	āhruron	*fell*	The Past 3rd Sg form of *āhrēosan* ‘to fall (down)’, a highly irregular Strong Verb (it does not survive into ModE).

31.	on	*on, onto*	> ModE *on*
32.	þæt	*the, that*	In OE, this is the Neuter form (SgNom or Acc) of *se*, so *þæt hūs* = 'the house' or 'that house'. The form survives into ModE as the demonstrative *that*, having – of course – lost its Neuter function.
33.	hit	*it*	> ModE *it*; has nothing to do with the verb *to hit*.
34.	nā	*(not) at all*	Originally a word meaning 'never', it is also used in OE to reinforce a negation; later on, it becomes a general negating word: ModE *no* derives from it.
35.	ne	*not*	The general negating particle in OE, always placed before the verb. It disappears during the Middle E period, being replaced by *not*.
36.	fēoll	*fell, collapsed*	The Past Sg 3rd person form of *feallan* 'to fall, to collapse', a Strong Verb (> ModE *fall/fell*). Altogether, *hit nā ne fēoll* = 'it did not fall at all'.
37.	sōþlīče	*truly, indeed*	An adverb formed from the adjective *sōþ*, meaning 'true'. It does not survive as an independent word, but cf. ModE *soothsayer*, i.e. 'some who tells the truth'.
38.	wæs	*was*	The 3rd person Sg Past Tense form of *be*, > ModE *was*.
39.	ğetimbrod	*built*	The Past Participle of *(ğe)timbrian* 'to build'. Cf. also the comments in row 21; *wæs ğetimbrod* = 'was built'.

5.3.2.2 A summary of the chief properties of Old English

1. Phonology

As you have seen, the fricative pairs **f/v, θ/ð, s/z** stood in complementary distribution: the voiced ones occurred inside a word between voiced sounds. This often resulted in alternations. The noun *wulf* 'wolf' for example, was pronounced with a **f** (= **wulf**) at the end, but the plural form *wulfas* had a **v** (= **wulvas**): the **f** in the plural form came to stand between voiced sounds so it automatically changed to **v**. While this voiceless/voiced alternation is no longer a regular feature of English, it still survives in some (by now irregular) relic forms, such as *wolves*, but also *loaf – loaves*, *knife – knives*. For a survival of the **θ/ð** alternation, cf. *path – paths*, *mouth – mouths*; as for **s/z**, cf. *house – houses*. Sometimes noun/verb pairs show the same pattern, cf. *half – to halve*, *bath – to bathe*, *house – to house*, where the verbs have a voiced fricative at the end. This is because in OE, the verbs (more precisely, the Infinitive) ended in *-an/-ian*, so the root-final fricative stood between voiced sounds. By ModE, the suffix was lost, but the original voiced fricative remained.

As noted above, word stress usually fell on the first syllable, except in certain prefixed forms in which the prefix was unstressed. By contrast, word stress in ModE is much more varied; the reason for this will be discussed in the next chapter.

Finally, there are many other interesting phonological phenomena in OE, the detailed discussion of which would be, however, beyond the scope of this textbook. See the Suggested Reading section if you are interested.

2. Morphology

Nouns, adjectives, and most pronouns still distinguish gender, viz. Masculine, Feminine, and Neuter. Adjectives have – as in PGmc - two types of inflection, Definite and Indefinite. If you check the text and the explanations above, you can see that nouns have various ways of forming the Plural, which I also illustrate (with additional examples) in (81) below.

(81) Some ways of forming the Plural in OE[41]

Singular (Nominative)	**Plural (Nom/Acc)**	**Meaning + OE Gender**
nam**a**	nam**an**	'name/names' Masc.
wind	wind**as**	'wind/winds' Masc.
sun**u**	sun**a**	'son/sons' Masc.
scip (pron. **ʃip**)	scip**u**	'ship/ships' Neut.
word	word	'word/words' Neut.
tal**u**	tal**a**	'tale/tales' Fem.
sunn**e**	sunn**an**	'sun/suns' Fem.
fōt	fēt	'foot/feet' Masc.
mūs	mȳs	'mouse/mice' Fem.

As you can see, the ways of forming the Plural were quite varied: it depended partly on Gender (e.g. the ending *–as* was found with Masculines only), but nouns of the same gender could have different plural forms, cf. *scip* and *word*; on the other hand, some endings (such as *–an*) were found in different genders.

The words in the shaded rows express the Plural by a vowel difference without a suffix. They are already highly irregular in OE, too, but some of them survive (as irregulars) up to the present, cf. also *man/men, louse/lice*, etc. In OE, there were more of them, such as *frēond/frīend* 'friend/friends', which have been regularized by now. Though the vowel difference reminds us of Ablaut, it has nothing to do with it – instead, it is a Pre-OE development (while Ablaut is inherited from PIE). Note that the Plural forms *fēt* and *mȳs* have a front vowel, while the Singular ones have a back vowel. In Pre-OE, the Plural forms originally had a suffix *–i*, added to the Singular form with a back vowel, so they were *fōti* and *mūsi*, respectively. Due to a sound change caused by the front vowel **i**, the back vowel of the root became front. This sound change is called by the (somewhat confusing) name **Umlaut** – quite similar to Ablaut, but not to be confused with it! Later on, the *–i* was dropped, and the vowel difference remained as the only indicator of Plural. To sum up, using the word *foot* as an example:

(82)	Singular	Plural
Original:	fōt	fōt-i
After Umlaut:	fōt	fēt-i
After the loss of final *–i*:	fōt	fēt

[41] More precisely, the Nominative/Accusative Plural (these two cases always had the same form in the Plural of nouns). The Genitive and the Dative forms are different, but they do not survive into ModE, so I leave them out here.

In the Singular, where there is no **i**, there is nothing to cause Umlaut, so the original back vowel remains.

Many Neuter nouns – such as *word* – have a suffixless or zero Plural, without a difference in the vowel of the root, so the Plural sounds the same as the Singular. In ModE, there are one or two relic forms, such as *sheep* (Plural: *sheep*).

The OE Plural form in *–an* (as in *naman, sunnan*) survives, for example, in the Plural of *ox*, viz. *oxen* (< OE **oxan**). It is also found in the irregular Plural forms *brethren* and *children.*

However, the most frequent Plural ending was *–as*, as in *windas*: indeed, about 30% of OE nouns had this Plural ending. In time, it was extended to other nouns, too, and by Middle English, it became the regular Plural suffix, replacing the other suffixes. I hope you have found out by now that the modern regular Plural suffix *-(e)s* originates from OE *–as*!

As for verbs, the essential properties have been mentioned in connection with PGmc. Some further notes:

As we have seen, negation is expressed by *ne* before the verb, as in *ne fēoll* 'did not fall'. Negation with *not* emerges during the Middle English period, and, if you recall, negation with the Auxiliary do appears as late as EModE – and even there, it is not obligatory.

An interesting point to note (although not clear from the text) is that most Strong Verbs had two different Ablaut forms in the Past Tense: one of them characterized the 1st and 3rd person Singular, the other the rest. The Past Tense of *singan* 'to sing', for example, was *sang* in the 1st/3rd person Singular, *sungon* in the Plural. By Modern English, Strong Verbs have given up this distinction (so you have *sang* in all persons), with one exception: *be*, which still shows the difference, cf. *I/he/she/it was* but *we/you/they were*. (There is also a difference in the final consonant, but it need not concern us here.)

3. Syntax

The most important difference between Old and Modern English is that word order is generally much more free in OE. This is because OE is more heavily inflected than ModE, and grammatical functions such as subject and object, for example, are mostly shown by different case forms. In Modern English, these functions are generally indicated by word order: *The king ate the dragon* does not mean the same as *The dragon ate the king*. In OE, however, they would differ in other ways than word order, too:

(83)

(a)	Se	cyning	æt	þone	dracan.
	the-Nom	king-Nom	ate	the-Acc	dragon-Acc
	'The king ate the dragon'				
(b)	Se	draca	æt	þone	cyning.
	the-Nom	dragon-Nom	ate	the-Acc	king-Acc
	'The dragon ate the king'				

As you can see, 'the dragon' has a different Nominative and Accusative form; and while *cyning* can be either, the form of the article – *se* vs. *þone* shows if *cyning* is a subject or an object. In fact, if I say *Þone cyning æt se draca*, it would still mean 'The dragon ate the king'!

This does not imply, of course, that word order is completely free in OE; it simply means that it is less strict than in Modern English. Indeed, it would be impossible to say

**draca se* instead of *se draca*, i.e. an article always comes before the noun, as in Modern English, too. Furthermore, there were some preferred word orders in some cases.

If you look at the following sentence, taken from the Matthew passage above, you find a word order which is not found in ModE:

(84)	hit	wæs	ofer stān	ğetimbrod
	it	was	on stone	built

Today, we would say *it was built on stone*, i.e. the Past Participle would come after the Auxiliary. In OE, this would have been possible, too, but there was another option, shown in (84): the Participle could also be placed at the end. This pattern, characteristic of Germanic in general, and still found in Modern German, is known as **clausal brace**, because the Auxiliary and the Participle "embrace" the rest of the clause.

4. Vocabulary

The overwhelming majority of OE words is of PGmc origin, and there are relatively few borrowings, most of which are (not too surprisingly) from Latin. These include words such as (given in their ModE forms) *chest, offer, fork, monk*, etc. In Mercia and Northumbria, invaded by the Vikings, many words were borrowed from North Germanic (Scandinavian), but they do not yet show up in West Saxon (which had never been conquered by the Vikings); since most of our texts are in West Saxon, we cannot tell for sure which words were borrowed into the Anglian dialects during the OE period. By Middle English, however, many words of Scandinavian origin become widespread over the whole of England; we will mention them in the next chapter.

Instead of borrowing, OE preferred word formation, using domestic elements, to express new concepts. Whereas ModE uses the word *trinity* (from Latin), for example, Old English used *þrines*, from the morphemes *þri* and *nes*, i.e. 'threeness'. Similarly, whereas we use the word *entrance* today (from French), OE used *ingang*, literally 'in-going' (cf. also Modern German *Eingang*). There are, of course, hundreds of other examples.

This concludes our discussion of Old English; let us now turn to our final chapter, looking at Middle English as well as some of the changes that took place between Middle and Early Modern English.

Suggested reading

For both PGmc and OE, check Part I of the Bibliography. Apart from these items, see also Volume I of *CHEL*. Robinson (1992) provides an introductory overview of the old Germanic languages. As for OE, there are several introductory textbooks, such as Hogg (2002) or Mitchell and Robinson (2012); Lass (1994) considers the historical background of OE, including many details on PIE and PGmc, too. As for detailed OE grammars, the classic one is Campbell (1959), but Hogg (1992) as well as its continuation, Hogg and Fulk (2011) are much more recent works.

6 Middle English and after

6.1 External history

As mentioned earlier, and as well known, the year 1066 marks a fundamental point in English history: this is the year when William, Duke of Normandy, claiming the English throne, landed with his army in the South of England, and on October 14th, he defeated the English army led by King Harold of England in the **Battle of Hastings**. He was crowned King of England at Christmas in the same year. This series of events is known as the **Norman Conquest**, and after that, William came to be known as **William the Conqueror**.

The Norman Conquest had serious – and devastating – consequences for Anglo-Saxon culture and literacy, as well as for society. Within two decades, William had replaced almost the entire English aristocracy as well as the higher clergy with his own men, who were, of course, from Normandy. Since the Normans spoke French (more precisely, the dialect known as **Norman French**)[42], this meant that the higher layers of society became French-speaking – while the common folk remained, of course, English-speaking. English – extensively used in the land before the Conquest – was ousted from noble circles and the court. The flourishing literacy of Anglo-Saxon England came to an abrupt end, and for some time, English was rarely used in a written form. The highly developed English standard, based on the West Saxon dialect, ceased to be used. It took several centuries – practically the whole of the Middle English period – until English gradually regained its prestige and came to be used again as an official language.

It would be a mistake to think, however, that the English language came close to extinction: it was still the language of the majority, and right from the outset, it did have some influence on the language of the Norman invaders. Indeed, the variant of Norman French spoken in England is referred to as **Anglo-Norman**, because it has some features which distinguish it from the French of Normandy. By about 1300, the Norman nobility and clergy had actually become English-speaking, adopting the language of the majority – although French was still learnt by aristocrats, but as a foreign language rather than as a native one; in other words, Anglo-Norman became extinct.

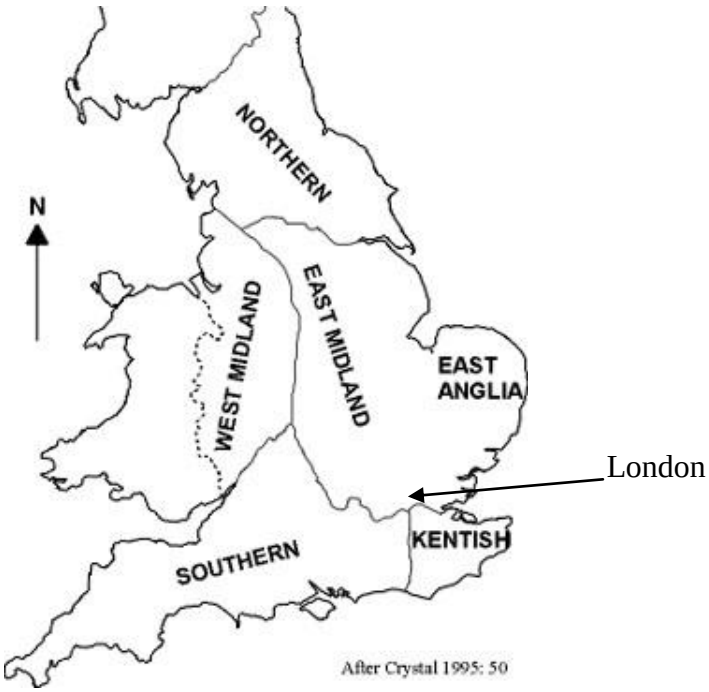
As far as the English language is concerned, it was also influenced by French, and quite significantly, in fact. The most important aspect of French influence was on vocabulary: English borrowed a huge amount of words from Anglo-Norman, but even after Anglo-Norman had ceased to be spoken, French influence on English vocabulary did not cease, although the source of French words was now **Central French**, the dialect of Paris and the surrounding area, which had a high prestige. Indeed, French continued to be a source of a number of words well into the Modern English period. The other important effect of French on English was in pronunciation, especially word stress. Both vocabulary and pronunciation will be examined later on.

As mentioned above, Anglo-Saxon literacy basically ceased to exist after the Norman Conquest. When English was used in writing at all, the author of the text used his/her (although it was mostly *he* at the time) own dialect, since there was no standard English. A result of this is that Middle English texts show significant dialectal variation: while Geoffrey Chaucer wrote in his own (London) dialect in the late 14th century, the author of *Sir Gawain and the Green Knight*, a contemporary of Chaucer, wrote in his West Midlands dialect, which

[42] Originally, the Normans were Vikings, speaking North Germanic. Soon after they invaded Normandy, however, they gave up their language, being assimilated to the French speaking population of the area. By the time of the Norman Conquest, they had become completely French-speaking.

was quite different from that of London. It is only towards the very end of the Middle English period that the dialect of London becomes quite firmly established as a standard. The chief dialects of Middle English (from now on: ME) are as follows: **Northern**, the descendant of the OE dialect called Northumbrian; **Kentish** (in the same area, i.e. Kent, as in OE times); **Southern**, corresponding roughly to the West Saxon area in the OE period; and finally, the Mercian dialectal area of OE had split into two different dialects, viz. **East Midlands** and **West Midlands**. You can see the ME dialects on the map below:

(85) ME dialects[43]



You can see that London is on the Southern periphery of the East Midland area, but very close to the Kentish and Southern dialects. This means that – although basically an East Midland dialect – the language of London shows several Southern and Kentish elements, too, from the earliest times on.

As far as ME literacy is concerned, then, we have noted two things. First, it is very varied dialectally speaking. Second, in the earlier part of the ME period, English was rarely used in writing, which is why we have relatively few surviving texts from Early Middle English; beginning with the 13th century, however, the number of texts starts to increase, and we have a range of texts from the 14th and 15th centuries. It would be impossible to list all of the most important works and authors; let me but mention the greatest author of the ME period, Geoffrey Chaucer, whose *Canterbury Tales* is an internationally known masterpiece (but he has written many other valuable works, too).

A further, third, thing to note is that when English started to be written again, due to the influence of French writing habits, several OE writing conventions had been abandoned, being replaced by French – or generally speaking, continental conventions. OE letters alien to continental writing fell into disuse: the letter *þ* (and the equivalent *ð*) was replaced by *th*, for instance. OE used the letter *c* for **k** as well as **tʃ**; in the latter function, it was replaced by *ch* (so, for instance, *cest* was respelt *chest* – no difference in pronunciation!). Some letter

[43] Source: http://www.yorkshiredialect.com/megen.htm, accessed 22/10/2012. Note: East Anglia is a region within the East Midland dialectal area.

combinations were also changed: word-initial **hw** – spelt *hw* in OE – came to be spelt as *wh* (as in *what*, cf. OE *hwæt*), and initial **kw**, spelt *cw* in OE, was replaced by *qu*, cf. *queen*, OE *cwēn*; in neither case is there any change in pronunciation. The letter *k*, not used in OE, was introduced, cf. *kin* vs. OE *cynn*. There are several other such spelling changes; the point is that several of the spelling conventions of ModE – indeed, the most basic ones – were born during the ME period, to a great extent due to French influence. Note that most of these spelling changes did not reflect any difference in pronunciation, that is, they are to be regarded as changes in written culture rather than the language itself.

This discussion leads us to our next topic: the pronunciation of ME.

6.2 Middle English pronunciation and phonology

Since most spelling conventions of ModE were established during the ME period, I give a rather detailed presentation of letter-to-sound rules – at least as far as vowels are concerned; consonant letters are mostly used in the same way as today (I will note the exceptions).

(86) ME vowel sounds and spellings[44]

	Sound	**Spelling**	**Examples (with ModE equivalents)**
1.	[a]	a	*sak (sack), shal (shall)*
2.	[aː]	a (aa)	*tale (tale), aker (acre), caas (case)*
3.	[e]	e	*bed (bed), seven (seven), herte (heart)*
4.	[ɛː]	e, ee	*clene (clean), deel (deal), breeth (breath)*
5.	[eː]	e, ee	*gees (geese), seke (seek)*
6.	[ɪ]	i, y	*hil (hill), kyn (kin)*
7.	[iː]	i, y	*hiden (hide), mys (mice)*
8.	[o]	o	*frogge (frog), cok (cock)*
9.	[ɔː]	o, oo	*hoom (home), boot (boat), throte (throat)*
10.	[oː]	o, oo	*goos (goose), mone (moon), good (good)*
11.	[ʊ]	u, o	*ful (full), bukke (buck), sonne (sun), comen (come)*
12.	[uː]	ou, ow	*mous (mouse), brow (brow)*
13.	[ə]	e	*tal<u>e</u> (tale), root<u>e</u> (root), all<u>e</u> (all)*
14.	[ɛɪ]	ai, ay, ei, ey	*hail (hail), day (day), seil (sail)*
15.	[ɔɪ]	oi, oy	*joy (joy), cloistre (cloister), joinen (join)*
16.	[aʊ]	au, aw	*lawe (law), laughen (laugh)*
17.	[ɪʊ]	eu, ew, u	*newe (new), vertu (vertue), fewe (few)*
18.	[oʊ]	ou, ow	*snow (snow), dough (dough), broughte (brought)*

[44] There were two more diphthongs, which, however, were very rare, and by Late ME, they fell together with other diphthongs, so I do not list them here.

Let us draw some important conclusions.

1. As you can see, ME – like ModE – has monophthongs (rows 1-13 in the table) and diphthongs (rows 14-18). Monophthongs are either short or long: each short monophthong has a long counterpart, except the schwa (ə), which is always short, and (as in ModE) always unstressed. Two of the short monophthongs, **e** and **o** have two corresponding long monophthongs, shown in the shaded rows; note that the spelling does not make a difference. Altogether, the 5 stressed short monophthongs correspond to 7 long monophthongs.

2. Note that the length of monophthongs *can* be indicated by doubling the letter (e.g. *aa, oo, ee*), but this is not consistent, except for the long **uː**, which is always distinguished from short **ʊ** (cf. rows 11 and 12). On the other hand, **ɪ** and **iː** are never distinguished in spelling, both spelt simple *i/y* (cf. rows 6 and 7).

3. The spelling of diphthongs is generally unproblematic, except that **ɪʊ** can be spelt *u*, which can also stand for short **ʊ** (cf. rows 11 and 17), and **oʊ** is spelt the same as **uː** (cf. rows 12 and 18).

As for consonants, note that *gh*, which is normally silent in ModE (e.g. *night, bough*, etc.) or is pronounced **f** in a couple of words (e.g. *laugh, cough*), is always pronounced as a voiceless velar fricative **x** in ME, so *night* is **niːxt**, *laugh* is **laʊx**, for instance.

The other important point is that all written consonant letters are pronounced, even when they are silent today, as in *know*, *write*, with **kn** and **wr**, respectively.

Finally, double consonant letters (as in *frogge* and *bukke* in the table) are pronounced long in ME (just like in Hungarian).

6.3 A sample text in ME with explanations

Let us now see a short sample text in ME, a verse of Chaucer's *Troilus and Criseyde*, followed by a phonetic transcription and a quasi-literal translation; then, some explanations are presented.

(87) The text in ME	ModE meanings of underlined words
Ye knowe eek that in forme of speche is chaunge	*also*
Withinne a thousand yeer, and wordes tho	*then*
That hadden pris, now wonder nice and straunge	*sense; very; odd, strange*
Us thenketh hem, and yet thei spak hem so;	*they seem to us; spoke; them*
And spedde as wel in love as men now do;	*succeeded*
Eek for to winnen love in sondry ages,	*various, different*
In sondry londes, sondry ben usages.	*lands; are; customs, habits*

(88) A phonetic transcription

jeː knoʊ eːk ðat ɪn form of speːtʃ ɪs ˈtʃaʊndʒə
wɪˈðɪnn a ˈθuːzand jɛːr, and ˈwoːrdəs θɔː
ðat ˈhaddən priːs, nuː ˈwʊndər niːs and ˈstraʊndʒə
ʊs ˈθeŋkəθ hem, and jet ðɛɪ spaːk hem sɔː
and spedd as wel ɪn lʊv as men nuː doː
eːk for toː ˈwɪnnən lʊv ɪn ˈsʊndrɪ ˈaːdʒəs
ɪn ˈsʊndrɪ ˈlɔːndəs, ˈsʊndrɪ beːn ɪʊˈzaːdʒəs

(89) A close translation

You know, too, that in the form of speech there is change
Within a thousand years, and words
That had a sense then, now seem to us very odd and strange
and yet they spoke them that way;
And they succeeded in love as well as people do now;
To win love, too, customs are different
in different ages and in different lands.

Let us see some notes now.
1. As you can see, several words of ME still exist in PdE, but with a different meaning, e.g. *nice*, *pris* (ModE *price*), *usages*.
2. In some cases, the word still exists, with the same meaning, but is archaic or rare, e.g. *sondry* (ModE *sundry*), or the verb *to speed* in the sense of 'to succeed'.
3. Some words have been lost, e.g. *eek* 'also', *tho* 'then', also *ben*, the typical Present Plural form of *be* in Chaucer's dialect, being replaced by *are*.
4. Two verb forms end in *–en*. The first one is *hadden* 'had', where it is a Plural suffix (*wordes ... hadden pris*, i.e. *words ... had sense*: *wordes* is a Plural subject, so you have *hadden*, rather than *had*). This is one of the functions of *–en* in ME. The other function is illustrated by *to winnen* 'to win': *-en* is also the suffix of the Infinitive (remember OE *–an*?). In both functions, however, it can be dropped already, cf. *ye knowe* (I hope you still remember *ye* from our discussion of EModE!), where the verb is in the Plural, yet it lacks the old ending (then, it would be *ye knowen*). Other instances are *thei spak* (instead of *thei spaken*), *(thei) spedde* (for *thei spedden*), and *men ... do* (for *men ... don*). The presence or absence of the suffix is optional in Chaucer's time.
5. Note the Plural noun form *yeer* 'years': this word was a Zero Plural noun, just like ModE *sheep* (cf. also the discussion of Plurals in OE).
6. The structure *us thenketh hem* – literally 'to us (it) seems them' – is an impersonal construction, whereas today we would say 'they seem to us'; note also the verb *thenketh*, by Chaucer's time a variant of *thinketh* 'thinks', which – besides its present-day meaning – also means 'seems'.
7. The word *hem* is an old pronoun form, meaning 'them'. It survives as the non-standard form *'em*, as in *I love 'em*.
8. Note the use of *for* before *to winnen*: it expresses purpose, i.e. *for to winnen* means '(in order) to win'. Such a structure is not possible any more.

6.4 Some important grammatical features of ME

Generally speaking, ME is definitely closer to Modern English – especially, of course, EModE – than OE. This is understandable, since it is closer to us in time. Besides, the fact that the basic writing conventions of English originate from ME makes the written language look more familiar. But the grammar is similar, too (although there are several differences). Let me mention some features which come into existence during the ME period.
1. The indefinite article *a/an* appears, whereas in OE, there was no indefinite article.
2. The definite article appears invariably as *the*, while in OE, it was inflected.
3. Word order is much closer to ModE: some characteristic word order patterns of OE, such as the clausal brace, have practically disappeared.

4. Old English had two tenses: Simple Present and Past. Perfect and Progressive forms, including the combination of the two, appear in ME: structures like *I have written, she is going, they have been waiting* are already found in ME, though they are relatively rare and are never obligatory: simple tenses can always be used instead (e.g. *she is going* means the same as *she goes* in ME).
5. In OE, the Comparative and the Superlative of adjectives was always formed with a suffix (as in ModE *nicer, nicest*). Comparison with *more* and *most* appears during the ME period, which may have been a result of French influence. Comparison with suffixes is still possible, however, in many cases where it would be impossible today, e.g. *cunningest* 'most cunning'.
6. The inflection of Adjectives is practically gone by ME, together with the definite/indefinite distinction. There is one remnant of the OE Adjective inflection: Adjectives consisting of one syllable may take an ending *–e* after determiners and in the Plural, *smal beest* 'small animal', vs. *the smale beest* 'the small animal' or *smale beestes* 'small animals'.
7. There are still many more strong verbs than in Modern English, but definitely fewer than in OE, a number of originally strong verbs having become weak. (Remember that regular verbs are weak, so this process can be regarded as an instance of regularization.)
8. Of course, the archaic features of EModE grammar – such as the use of a distinct 2nd person Singular personal prounoun (= *thou*) as well as the use of the verbal suffix *–est* (as in *thou comest*) and the 3rd Sg *–th* (as in *she loveth*) – are also characteristic of ME.

Needless to say, this short list does not aim at an exhaustive description of ME grammar. If you are interested in it in more detail, please check the items in the Suggested Reading section.

6.5 The vocabulary of ME and word stress in borrowings

6.5.1 Borrowings in ME

As mentioned earlier, the ME period sees the influx of a huge number of loanwords. The main sources are **Scandinavian**, **French** and **Latin**.
1. We mentioned in the previous chapter that most of England lived under Scandinavian (Viking) rule for a long time. The co-existence of English and North Germanic in the areas outside Wessex resulted in substantial Scandinavian influence on English. However, since before the Norman Conquest, the West Saxon dialect (unaffected by Scandinavian influence) was used as a standard, words of Scandinavian origin did not appear in writing. After the Norman Conquest, however, West Saxon ceased to be used as a standard, and people wrote in their own dialect. As a result, Scandinavian words – which had been there in spoken English well before the Conquest – started to appear in written texts.

The striking fact about Scandinavian loanwords in English is that there are many common words among them, some of which belong to the basic word stock, which tends to resist borrowing. The reason why basic words could be easily borrowed was that Old Norse, the language of the Vikings, was still very close to Old English, so much so that the English and the Vikings could even understand each other's speech. Basically, they can be regarded as dialects, rather than really different languages, for practical purposes. In such a situation, borrowing is very easy. Let me give a couple of words of Scandinavian origin: *take, sky, skin, skirt, law, give, get, sister*, but even the pronouns *they/them/there are* from Scandinavian! In many cases, the Scandinavian form simply replaced the similar OE form, as in the case of *give* and *sister*; in other cases, the word is a total newcomer, like the verb *take* (the OE word, *niman* 'to take' was not similar at all). Sometimes the word enters the English language and

the OE word remains, too, as in the case of *skirt*, whose OE cognate, *shirt*, also survives, but the two have different (though clearly related) meanings.
2. French loans (as mentioned) start to arrive after the Norman Conquest. Since French never became the language of the common folk, most words of French origin are (at least originally) not basic, everyday ones, but cultural, political, legal, etc. terms. Examples are numerous, e.g. *war, castle, parliament, prison, service, justice, punish,* to mention but a few. In time, however, more common words were also borrowed, such as *large, question, city, chair, dance, chance, change, catch, finish, age, number, chief,* to mention but a few. Sometimes the French word co-exists with a native (OE) one, but their meaning is different: *city* vs. *town*, *chair* vs. *stool* are fine examples. Sometimes a common word is borrowed but with a specialized meaning: the word *noun*, for instance, simply meant 'name' in Anglo-Norman! But probably the funniest ones are those pairs where the native English word means a large edible animal – and the French loanword refers to its meat! Examples include *pig – pork, cow – beef, calf – veal, sheep – mutton*. The anecdotal explanation is that the French nobility ate these animals, while the common English folk only saw them, but did not eat them!
3. Latin loans had already existed in OE, but their number started to increase rapidly after the Norman Conquest; it would be a hopeless enterprise to give a representative sample, since the Latin loans in English probably number thousands, if not tens of thousands. During the ME period, we have words such as *testament, discuss, complete, allegory, minor, necessary, equal, private, quiet*, etc. It must be added that many Latin words arrived via French, and it is often difficult or impossible to tell if a given word is from Latin or from French.

6.5.2 Word stress in borrowings

Recall that words were generally stressed on the first syllable in OE: indeed, this pattern is found in all old Germanic languages. The loanwords of Scandinavian origin, therefore, did not differ: the Vikings also spoke Germanic, so they, too, stressed the first syllable of words. The situation is quite different in the case of French loans, since Old French (Anglo-Norman and Central French alike) stressed words towards the end, either on the last syllable or the last but one. Interestingly, ME borrowed this tress together with the words rather than assimilating them to the initial stress pattern of Germanic. Words with French-type stress (given in their ME form, followed by a phonetic transcription and the ModE form) include *nacióun* (= **ˌnaːsiˈuːn** 'nation'), *vanité* (= **ˌvaniˈteː** 'vanity'), *presént* (noun; = **preˈzent** 'présent'), *preséntenen* (verb; **preˈzentən** 'presént'), *cité* (= **sɪˈteː** 'city'), *adversité* (= **adˌversiˈteː** 'adversity'), etc. As you can see, word with three or more syllables had a secondary stress on the third-last syllable. Also, by ModE, the stress moved the first syllable of disyllabic words in the majority of cases (cf. ModE *cíty, présent*$_N$) but not in prefixed verbs (cf. ModE *preséntV*): in fact, recall that prefixes were unstressed in OE, too, cf. prefixed verbs of OE origin such as *becóme, aríse*, etc. Words with more than two syllables moved the primary stress to the third-last, originally secondary stressed syllable, cf. *vánity, nátion, advérsity*. Altogether, the stress tends to move towards the beginning of the word, making it closer to the Germanic stress pattern, but there are still many words which are not stressed on the first syllable. Latin borrowings were also generally stressed according to the Old French pattern. The different stressing of French/Latin as opposed to native words is the chief reason why word stress is so varied in Modern English.

6.6 Vowel length in ME

ME, just like OE, distinguished short and long vowels – the latter also including diphthongs. Furthermore, recall that the schwa was only found in unstressed syllables, and it was always short: we call it a **weak** vowel, as opposed to all the others, termed **full** vowels. Within the latter group, long and short monophthongs could distinguish words: *god*, for example, had a short **o**, while *good* was pronounced with a long **oː**. On the other hand, there were some cases where only a short vowel was possible, and conversely, there were cases where the vowel was generally long. There are two such regularities I would like to discuss here, because they have some relevance for two general spelling conventions of ModE.

1. In words of two syllables stressed on the first syllable, the stressed vowel was long if it was <u>followed by a single consonant and a word-final schwa</u>. An example is the word *name*, pronounced **ˈnaːmə** in ME: the stressed vowel is followed by a single consonant, itself followed by a word-final schwa. Other examples include *take, tale, mete (= meat), weke (= week), hope*, etc. These words all had a long stressed vowel. By Late ME, however, the final schwa was dropped, so *name*, for example, became **ˈnaːm**. In the spelling, the letter *–e* (now silent) remained, and it soon came to be interpreted as a way of showing the length of the stressed vowel. Accordingly, it soon started to be used as a marker of vowel length even in cases where there had not been a final schwa. For instance, the word *stone* is spelt with a silent *–e* at the end, although it had never ended in a schwa: in OE, it was *stān*, in ME, *stoon*, pronounced **stɔːn**. Silent *–e* still has this basic function: compare the word *sit*, with a short vowel (and no final *–e*), vs. *site*, where the final *–e* is there to show that the *i* is to be pronounced long, i.e. as **aɪ**, rather than **ɪ**. Further such pairs include *mad – made, hat – hate, met – mete, dun – dune, cod – code*, etc.

2. Contrary to the above situation, stressed vowel were short before certain combinations of consonants, including several obstruent clusters such as **pt, ft, kt**, etc. Crucially, double (long) consonants also counted as a cluster of two consonants (where the two are identical): before a double **tt**, for instance, ME had only short vowels, just like before **pt**. Let us see some examples[45]. The verb *keep* had a long vowel in ME (like in ModE): it was pronounced **keːp**. The Past Tense form, however, had a suffix *–te*, so it combined with the **p** to yield a form **keːptə**, which shortened to **keptə** because of the **pt** cluster. The Past Tense of *meet* (pron. **meːt**) was similar: a suffix *–te* was added to it, yielding **meːttə**, shortened to **mettə** because of the double **tt**. Some verbs ending in **d**, such as *feed, hide*, etc., also showed shortening: they had a suffix *–de* in the Past Tense, which then combined with the final **d** of the verb, producing a double **dd**: e.g. *feed*, pronounced **feːd**, added **də**, yielding **feːddə**, which in turn shortened to **feddə**. The table below provides a summary:

(90)

Present:	ME pron.:	Past (ME spelling)	ME pron.:	ModE spelling:
keep	**keːp**	*kepte*	**kep<u>t</u>ə**	*kept*
meet	**meːt**	*mette*	**me<u>t</u>tə**	*met*
feed	**feːd**	*fedde*	**fe<u>dd</u>ə**	*fed*
hide	**hiːd**	*hidde*	**hɪ<u>dd</u>ə**	*hid*

As for the Past forms are concerned, note that they had a schwa at the end, which – as mentioned above – was dropped in the later part of the ME period. As you have probably discovered, too, the forms with ME long **tt** and **dd** no longer have a long (double) consonant,

[45] I disregard any further suffix in this brief presentation, since it is not relevant for our point.

neither in pronunciation nor in spelling. The reason is that all long consonants were shortened by the Early Modern English period. This means by the beginning of the 16th century, the forms *mette, fedde, hidde*, etc. were pronounced just like today, i.e. **met, fed, hɪd**. (They were also soon respelt to reflect the new pronunciation.) Middle English, however, had long consonants in many other cases, too, not only in Past Tense forms; examples include *appel* (ModE *apple*), *better, sitting, spilling, croppes* (ModE *crops*), *sonne* (ModE *sun*), to mention but a few. After the long consonants (before which, remember, the vowel had always been short) shortened, an interesting – though not unexpected – thing happened as far as the relationship between pronunciation and spelling is concerned: the double consonant *letters*, which by now were not pronounced long any more, came to be reinterpreted as indicators of the *shortness of the preceding vowel*. As a result, the doubling of the consonant letter was extended to many other words with a short stressed vowel – words which had never had a long consonant! The word *summer*, for example, has never had a long **mm** (in OE, it was *sumor*, in ME, *sumer*); but it has a short vowel (ME **ʊ**, ModE **ʌ**), so the letter *m* was doubled in the spelt form to show that the vowel is pronounced short. The same thing happened in *written* (ME *writen*), *ridden* (ME *riden*), *hammer* (ME *hamer*), etc. The doubling of the final consonant letter in the *–ing* form of verbs, when the consonant is preceded by a short stressed vowel, has become a regular feature of ModE spelling, cf. *swim/swimming, pat/patting, compél/compélling*, etc., and you find the same doubling in regular Past Tense forms (e.g. *patted, compélled*, etc.). The doubling of consonant letters, however, has never been applied with full consistency: think of words like *lémon, Brítish, cíty, mánor*, etc., where the stressed vowel is short but the consonant letter after it is not doubled.

If you go back to (90), you can see that we have explained the Past Tense forms all right, but what about the Present forms? After all, they have quite different vowels today: *keep, meet, feed* have **iː**, rather than **eː**, and *hide* is pronounced with an **aɪ**, not with an **iː**. In fact, note that the Present/Past pairs have practically the same (or at least very similar) vowel: the only difference is in length. In ModE, on the other hand, this is not the only difference – the two forms have completely different vowels (e.g. *meet/met, hide/hid*). The reason for this is a wholesale transformation of the long monophthongs of English, which we are turning to in the next section.

6.7 Long monophthongs from ME to EModE: The Great Vowel Shift

The long monophthongs of ME undergo a radical chain shift during the late 15th and early 16th centuries. This is called the **Great Vowel Shift** (GVS). Being a chain shift, it is quite similar in the way it happens to Grimm's Law (but, of course, it affects vowels). The picture I give here is a simplified one (as it was with Grimm's Law, too), but it will give you a general idea of the change.

The table in (91) presents the long monophthongs of ME according to their articulation (on the left), with appropriate examples (on the right).

(91)

	Front	Back
High	**iː**	**uː**
Mid-high	**eː**	**oː**
Mid-low	**ɛː**	**ɔː**
Low	**aː**	

Examples:

	Front	Back
High	*time*	*house*
Mid-high	*meet*	*moon*
Mid-low	*peace*	*stone*
Low	*name*	

The word *time*, therefore, was pronounced with an **iː**, *house* with an **uː**, etc. Note that there is no low back long vowel, i.e. nothing below **ɔː** (this is shown by the shaded box in (91).

Step 1 of the GVS: high long monophthongs become diphthongs. Specifically, **iː** changes to **əɪ**, while **uː** becomes **əʊ**; *time* and *house* are now pronounced as **təɪm** and **həʊs**, respectively. (If you go back to Chapter 3, you can check the vowels of EModE!) In other words, the ME long high vowels disappear from the system of long monophthongs! This gives us the following situation:

(92) The situation after Step 1 of the GVS

	Front	Back
High		
Mid-high	**eː**	**oː**
Mid-low	**ɛː**	**ɔː**
Low	**aː**	

Examples:

	Front	Back
High		
Mid-high	*meet*	*moon*
Mid-low	*peace*	*stone*
Low	*name*	

There is no no-one in the shaded rows. Now comes Step 2 of the GVS: Mid-high vowels move to the empty places, becoming high: **eː** changes to **iː**, while **oː** becomes **uː**. This gives us the picture below:

(93) The situation after Step 2 of the GVS

	Front	Back
High	**iː**	**uː**
Mid-high		
Mid-low	**ɛː**	**ɔː**
Low	**aː**	

Examples:

	Front	Back
High	*meet*	*moon*
Mid-high		
Mid-low	*peace*	*stone*
Low	*name*	

The mid-high long vowels are now gone; the words *meet* and *moon* are pronounced with a high vowel, just like today. Step 3 of the GVS will now take place, whereby the mid-low long vowels move up to the empty rows, i.e. **ɛː** changes to **eː**, while **ɔː** becomes **oː**, shown in (94):

(94) The situation after Step 3 of the GVS

	Front	Back
High	**iː**	**uː**
Mid-high	**eː**	**oː**
Mid-low		
Low	**aː**	

Examples:

	Front	Back
High	*meet*	*moon*
Mid-high	*peace*	*stone*
Mid-low		
Low	*name*	

Now, the mid-low vowels have disappeared, *peace* and *stone* pronounced with a mid-high vowel, as **peːs** and **stoːn** (cf. Chapter 3). Finally, Step 4 of the GVS takes place: the long low front **aː** becomes mid-low, i.e. **ɛː**, but – since there is no vowel in the "low back" box, there, nothing happens. The end result is depicted in (95):

(95) The situation after Step 4 of the GVS

	Front	Back
High	**iː**	**uː**
Mid-high	**eː**	**oː**
Mid-low	**ɛː**	
Low		

Examples:

	Front	Back
High	*meet*	*moon*
Mid-high	*peace*	*stone*
Mid-low	*name*	
Low		

Although there are now 5, rather than 7, long monophthongs, the overall number of the original vowels has not changed: two new diphthongs have been created by Step 1 of the GVS!

In fact, due to other changes, the empty rows in the "back" column have also been re-filled: this is shown in our last table. Note, however, the two "newcomers" in the back column (shown by the symbol ☞) are ***not*** the results of the GVS, but of different changes.

(96) The long monophthongs of EModE

	Front	Back
High	**iː**	**uː**
Mid-high	**eː**	**oː**
Mid-low	**ɛː**	☞ **ɔː**
Low		☞ **ɑː**

Examples:

	Front	Back
High	*meet*	*moon*
Mid-high	*peace*	*stone*
Mid-low	*name*	☞ *law*
Low		☞ *palm*

Let us now go back to the problem mentioned at the end of Section 6.6; you can now understand why the vowels are different in *keep* vs. *kept*, or *hide* vs. *hid*. In the Present Tense form, there was a long monophthong, which underwent the GVS, but in the Past Tense form, the vowel was short – and short vowels were not affected by the GVS at all!

The GVS also helps to understand why vowel letters have – from a "continental" point of view – some strange pronunciations in English. There is no other European language, for instance, where the letter *i* sounds as **aɪ** (as in English *time*), or the letter *a* is pronounced **eɪ**, as in *name*. The trick is, very simply, that the GVS (not to mention later changes, discussed in Chapter 3) profoundly altered the long vowels, but the spelling has remained basically the same as it was in the early 15th century, that is, before the GVS. Indeed, it is generally true that the highly conservative nature of English writing is one of the chief reasons (if not the most important one) why the pronunciation of English words causes so many problems to foreign learners of English – and conversely, why learning to spell correctly takes so much time and energy for English-speaking schoolchildren, and why spelling competitions are so popular in English-speaking countries.

Suggested reading

First, please check Part I of the Bibliography as well as Volume II of *CHEL*. General detailed descriptions of ME include Horobin and Smith (2002), Burrow and Turville-Petre (2004) as well as Mossé (1952) – this last one is not very recent, but it is a good and thorough book. For changes that took place in Late ME and EModE, see the Suggested reading section in Chapter 2 and Chapter 3.

BIBLIOGRAPHY

Part I: Some general introductions to the history of English

Algeo, J. and T. Pyles (2005) The origins and development of the English language. 6th ed. Wadsworth and Cengage Learning.

Barber, C., J. C. Beal, and P. A. Shaw (2009) The English language: A historical introduction. Cambridge: Cambridge University Press.

Baugh, A. C. and T. Cable (2002) A history of the English language. 5th ed. London: Routledge.

Cser, A. (2003) An outline of the history of the English language. Piliscsaba: PPKE.

Hogg, R. M. and D. Denison (eds.) (2006) A history of the English language. Cambridge: Cambridge University Press.

Strang, B. M. H. (1970) A history of English. London: Methuen.

Part II: CHEL = The Cambridge History of the English language, published by Cambridge University Press, in 6 volumes:

Volume I (1992): The beginnings to 1066. Edited by Richard M. Hogg.

Volume II (1992): 1066-1476. Edited by Norman Blake.

Volume III (1999): 1476-1776. Edited by Roger Lass.

Volume IV (1998): 1776-1997. Edited by Suzanne Romaine.

Volume V (1994): English in Britain and overseas: Origins and development. Edited by Robert Burchfield.

Volume VI (2001): English in North America. Edited by John Algeo.

Part III: Other

Aitchison, J. (2001) Language change: Progress or decay? 3rd ed. Cambridge: Cambridge University Press.

Baldi, P. (1983) An introduction to the Indo-European languages. Carbondale, IL: Southern Illinois University Press.

Balogné, Bérces Katalin, and Szentgyörgyi Szilárd (2007) The pronunciation of English. Budapest: Bölcsész Konzorcium. Downloadable at: http://mek.oszk.hu/04900/04910/

Barber, C. (1997) Early modern English. Edinburgh: Edinburgh University Press.

Burrow, J. A. and T. Turville-Petre (2004). A book of Middle English. 3rd ed. Oxford: Blackwell.

Bynon, T. (1983) Historical linguistics. Cambridge: Cambridge University Press.

Campbell, A. (1959) Old English grammar. Oxford: Clarendon Press.

Campbell, L. (2004) Historical linguistics. 2nd ed. Cambridge, MA: MIT Press.

Clackson, J. (2007) Indo-European linguistics: An introduction. Cambridge: Cambridge University Press.

Fortson, B. W., IV. (2011) Indo-European language and culture: An introduction. 2nd ed. Oxford: Blackwell.

Görlach, M. (1991) Introduction to Early Modern English. Cambridge: Cambridge University Press.

Hogg, R. M. (1992) A grammar of Old English. Vol. I: Phonology. Oxford: Blackwell.

Hogg, R. M. (2002) An introduction to Old English. Edinburgh: Edinburgh University Press.

Hogg, R. M. and R.D. Fulk (2011) A grammar of Old English. Vol. II: Morphology. Oxford: Blackwell.

Horobin, S. and J. Smith (2002) An introduction to Middle English. Oxford: Oxford University Press.

Lass, R. (1994) Old English. Cambridge: Cambridge University Press.

Lehmann, W. P. (1993) Historical linguistics. 3rd ed. London and New York: Routledge.

McMahon, A. M. S. (1994) Understanding language change. Cambridge: Cambridge University Press.

Mitchell, B. and F. C. Robinson (2012) A guide to Old English. 8th ed. Oxford: Wiley-Blackwell.

Mossé, F. (1952) A handbook of Middle English. Baltimore: John Hopkins.

Nádasdy (2007) Practice book in English phonetics and phonology. Budapest: Nemzeti Tankönyvkiadó.

Robinson, O. W. (1992) Old English and its closest relatives. Stanford, CA: Stanford University Press.

Smith, J. (1999) Essentials of Early English. London and New York: Routledge.

Szemerényi, O. (1999) Introduction to Indo-European linguistics. Oxford: Oxford University Press.