Language change: the dynamicity of linguistic systems

KEY TERMS
- Synchrony and diachrony
- Sound change
- The Great Vowel Shift
- Grammaticalization
- Renewal
- Analogy
- Cognates
- Language families
- Proto-languages
- The Comparative Method
- Linguistic paleontology

CHAPTER PREVIEW
In previous chapters, we have talked about languages of the world and how they have different ways of expressing similar meanings. We have also examined various specific aspects of linguistic structure, including phonology, morphology, and syntax. But how did these structures within each language come to be? And how do languages come to be so different from each other?

This chapter describes how languages evolve over time: how sounds can shift, how sound systems can be restructured, and how grammar develops. It then discusses genetic relationships among languages and how they can be detected. It closes with a glimpse of what reconstructed languages can tell us about their speakers. An important aspect of language change is the fact that it is not arbitrary; it is the result of acts of speakers, as they learn their languages, use them, repair them, and extend them to use in new situations and for new purposes. The investigation of language change can tell us much about why languages are as they are, with all of their seeming irregularities. It can also provide us with a view of the human mind at work, recognizing patterns, and extending them to new situations and uses.

LIST OF AIMS
After reading this chapter, students should be able to:
- name and briefly describe some common kinds of sound changes;
- give examples of sound changes that have occurred in the development of English;
12.1 **Language change**

Earlier chapters have shown that languages are tightly patterned, in ways speakers rarely imagine. Because languages seem to operate so smoothly, we might assume they are inert; nothing could be further from the truth. *All living languages are dynamic, constantly being reshaped by their speakers.* Most of these adjustments are so gradual that they go unnoticed. Yet when we look carefully at the kinds of changes that occur in language, we can catch glimpses of the human mind at its most creative.

Apart from occasional new technical terms or slang expressions, we are rarely aware of language change. Can you identify the language below? Can you understand it?

\[
\text{Sigemunde gesprong æfter dēaðe dōm unlytel,} \\
\text{syðlan wiges heard wyrm ācwealde, hordes hyrde.}
\]

It is English, but English at an earlier stage of development. This Old English passage is from *Beowulf*, a manuscript written well over a thousand years ago.

(1) **Beowulf** 884–887

\[
\text{sigemunde gesprong æfter dēaðe dōm unlytel} \\
\text{to. Sig mund has arisen after death. day glory great} \\
\text{‘To Sigmund sprang forth great glory after his death,} \\
\text{syðlan wiges heard wyrm ācwealde hordes hyrde} \\
\text{Since of. fight brave serpent killed. off. of. treasure guardian} \\
\text{because in brave warfare, he killed the dragon, guardian of the treasure.’}
\]

Can you detect any similarities between the language in this passage and modern English? Some of the words are immediately familiar: æfter ‘after,’ dēað ‘death,’ and dēg ‘day.’ Some become identifiable after some consideration. Unlytel is ‘un-little’ or ‘not small’; heard is the ancestor of modern ‘hardy,’ hord of modern hoard, and hyrde of modern herd, as in shepherd. But there are certainly differences between the English of then and now. These differences fall under several types:

- **Phonological:** The noun dēg ‘day’ has a g at the end. The vowel in heard is not the same as that of modern hardy.

- **Morphological:** The verb ge-sprong has a prefix ge, an old marker of past participles (adjectival forms of verbs). Many of the nouns have case suffixes, which identify their
roles in the sentence. The nouns *wīg* ‘fight’ and *hord* ‘treasure,’ for example, end in -es, the genitive marker meaning ‘of.’ The name *Sigemund-e* ends in the dative marker -e meaning ‘to.’ The compound *dēaðdaeg-e* ‘death day’ also ends in the dative -e, which identifies it as the object of the preposition ‘after.’

- **Syntactic:** The verb ‘sprang’ occurs in second position in the first clause, before its subject ‘great glory.’ This word order is occasionally found in stylized constructions today, but it is rare in everyday conversation. In the second clause, the object ‘dragon’ occurs before the verb ‘killed,’ which is not a possible order today. Modern English word order is strongly SVO.
- **Lexical:** Some of the vocabulary has been replaced by other words in modern English, such as *wīg* ‘fight, warfare.’
- **Semantic:** The meanings of a number of the words have changed. *Dōm* ‘glory’ has developed into modern *doom*; *wyrm* ‘serpent, dragon’ into modern *worm*; *cweal-de* ‘kill-ed’ into modern *quell-ed*.

### 12.2 Sounds

Chapter 3 described ways that speakers alter their pronunciation of sounds. These alternations can result in allophones, discussed in Chapter 3. Over time, the same kinds of processes can result in significant sound change.

#### 12.2.1 Here and there: sound change

English spelling is notorious for being difficult to learn. It seems to fit the language badly. There is a reason for this. Old English was written in a version of the Roman alphabet introduced by Irish missionaries. Words were spelled according to their pronunciation at the time, with letters representing their Latin values. As sounds changed over the course of the development of English, spelling practices changed too, but not as quickly as pronunciation. Words now spelled with final *e*, such as *her(e)* and *ther(e)*, were once pronounced with a full final vowel. As time went by, this unstressed *e* underwent **lenition** (weakening). By Chaucer’s time in the late fourteenth century, it had become reduced to schwa [ǝ]. Further lenition resulted in complete **loss**.

Sound changes can be described with rules much like those used in phonology. The term **synchrony** is used to refer to language structures at a single point in time. Synchronic rules, such as those discussed in Chapter 3, are operative at a specific moment, typically the present. The term **diachrony** is used to refer to comparisons of two or more points in time. A **diachronic rule** describes a change in the system between one time and another. We could describe the lenition of the vowel *e* to schwa with a diachronic rule like the one below:

*\(e > \mathbf{\dot{e}}\)  (The vowel *e* became schwa.*)
The asterisk * marks a sound that we infer existed at an earlier time; we have not actually heard it, but we hypothesize what it might have been, in this case on the basis of early documents. Where synchronic phonological rules contain an arrow \( \rightarrow \), diachronic rules usually have the symbol \( \Rightarrow \), which can be read as ‘became.’ (See Sidebar 12.1 for a note on the differences between rule-writing in historical linguistics and phonology.)

The lenition of \( e \) to schwa that we saw in *here* and *there* did not occur everywhere in the language. It happened only at the ends of words: it was what is called a conditioned sound change. We can specify the context in which it occurred (only at the ends of words), in the same way as in a synchronic phonological rule. We add a slash / to announce that the environment for the change is coming up. We use an underline (a blank) __ to stand for the position of the sound in question. We then specify what sounds occur before it to the left of the blank, and what sounds occur after it to the right. In this case, it does not matter what occurred before the \( e \), so nothing appears on the left. The crucial context is what occurred after the \( e \), namely a word boundary. This is indicated by the symbol #.

\[
*e \Rightarrow \text{schwa} / \_\_ # \quad (The \ vowel \ e \ became \ schwa \ at \ the \ ends \ of \ words.)
\]

The weakening or leniting of word-final \( e \) did not stop at schwa. It continued over time until there was no vowel left at all. We can add this fact to our rule, with a zero:

\[
*e \Rightarrow \text{schwa} \Rightarrow \emptyset / \_\_ # \quad (The \ vowel \ e \ became \ schwa \ and \ was \ then \ lost \ at \ the \ ends \ of \ words.)
\]

There are numerous other examples of lenition that took place during the development from Old English to Modern English. As we saw in (1), the noun ‘day’ was spelled \( \text{dæg} \). The final \( g \) was pronounced at the time, but it later underwent lenition to a velar fricative, then was further weakened to just a glide. Sidebar 12.2 provides another example of the lenition of a velar fricative, still reflected in the English spelling system.

### SIDEBAR 12.2
**ENGLISH SPELLING**

Have you ever wondered why English words like ‘night’, ‘light’, and ‘right’ are spelled with \( gh \)? Those words once contained velar fricatives \( [x] \), which were first lenited (weakened) to \( [h] \), and then ultimately lost.

For further exploration of the many irregularities of English spelling, work through the examples in Sidebars 12.3 and 12.4.

12.2.2 Trees and leaves: change in the system

**Sound change can have more profound effects; it can result in the remodeling of the sound system.** The most common way to form a plural in present-day English is to add the suffix \(-s\) to the basic form of the noun: *tree/tree-s, flower/flower-s*. But not all words follow this pattern. Consider the plurals of *leaf, wife,*
and shelf: leaf/leave-s, wife/wive-s, and shelf/shelve-s. The noun roots change shape: /li:f/ /li:v/-, /waif/ /waiv/-, /ʃɛlf/ /ʃɛlv/-.

As seen in Chapter 3, phonemes may have alternate pronunciations, called allophones. A common process that produces allophones is assimilation, whereby a sound comes to be more like its neighbors. Old English originally had just one series of fricative phonemes: /f/, /θ/, /s/, /ʃ/, and /x/. Each of these fricatives had allophones. When a fricative occurred between voiced sounds, it was automatically voiced too. We could write a synchronic phonological rule to describe the allophony at that time.

\[
\text{[fricative]} \rightarrow \text{[+voice]} / \text{[+voice]} \_ \text{[+voice]}
\]

Fricatives assimilated to their neighbors in voicing.

The voicing was predictable: the fricative [v], for example, occurred only inside of words, between voiced sounds, while [f] occurred everywhere else, that is, at the beginnings of words, at the ends, or next to a voiceless sound. This is a familiar pattern of complementary distribution, discussed in Chapter 3.

With the Norman invasion in 1066, French speakers began to arrive in England in significant numbers. As time went by, the influence of their language on English became increasingly apparent. Many French words were adopted into English. The adverb very, for example, came into English from the Old French *vrai* (modern French 'true'). The ancestors of the present-day English words veal, veil, venerable, vengeance, venison, venom, vent, vermin, vessel, villain, vine, voyage, and many more were borrowed from French. The influx of so many words beginning with *v* upset the old pattern of complementary distribution. There were now minimal pairs such as fine and vine. Voicing was no longer predictable. The result was a change in the phonological system: the addition of a new phoneme /v/. **Sounds that were originally allophones became separate phonemes.**

12.2.3 The moon and the goose: the Great Vowel Shift

Have you ever wondered why the letter *o* is pronounced [o] in words like so, but [u] when it is doubled, as in *moon* and *goose*? Vowels are some of the most puzzling aspects of the English spelling system.
As mentioned earlier, when the Roman alphabet was adopted for the spelling of Old English, the vowel letters were used to represent approximately the same sounds as in Latin. Like Latin, Old English had both long and short vowels: long vowels simply lasted longer than short ones, as in *moon* and *keep*. At a certain point, English speakers began to shorten long vowels in certain contexts, a conditioned sound change. One of these contexts was before consonant clusters, as in *monthly* and *kept*. Then, beginning around 1400, vowel sounds began to shift. Long vowels began to rise. Long low vowels (ā, ĺ̄) were pronounced as mid (ǣ, ō). (English vowel length is traditionally written with a macron, or a bar, over the vowel.) Mid vowels (ē, ō) were pronounced as high (ī, ū). The highest long vowels (ī, ū), which could not rise any higher, broke into diphthongs (ōi, ūa). These changes, which took place gradually over a period of three centuries, are referred to collectively as the Great Vowel Shift. Stages in the process can be seen in Table 12.1.

Because of these shifts, the pronunciation of vowels in Modern English is now very different from that in Old English, when the spelling system originated. This history explains some of the puzzles of modern spelling. The words *moon* and *goose* are spelled with *oo* because they were once pronounced with long *o*. The words *beet* and *beat* are spelled differently because they were once pronounced differently: there was a merger. The two vowels ĭ and ē merged by the time of Wordsworth to ĭ. The result was a change in the system, from seven long vowels (including diphthongs) to six, and a reduction in the number of long vowels that were not diphthongs.

### Table 12.1 The Great Vowel Shift (Anttila 1972: 65)

<table>
<thead>
<tr>
<th></th>
<th>Chaucer (1343–1400)</th>
<th>Shakespeare (1564–1616)</th>
<th>Wordsworth (1770–1850)</th>
<th>Modern English (present)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>bite</em></td>
<td>ĭ</td>
<td>ĭi</td>
<td>ĭi</td>
<td>ai</td>
</tr>
<tr>
<td><em>bete</em></td>
<td>ēi</td>
<td>ĭi</td>
<td>ĭi</td>
<td>ĭi</td>
</tr>
<tr>
<td><em>bete</em></td>
<td>ēi</td>
<td>ĭi</td>
<td>ĭi</td>
<td>ĭi</td>
</tr>
<tr>
<td><em>name</em></td>
<td>āi</td>
<td>āi</td>
<td>āi</td>
<td>āi</td>
</tr>
<tr>
<td><em>foal</em></td>
<td>ōi</td>
<td>ōi</td>
<td>ōi</td>
<td>ou</td>
</tr>
<tr>
<td><em>fol</em></td>
<td>ōi</td>
<td>ōi</td>
<td>ōi</td>
<td>ōi</td>
</tr>
<tr>
<td><em>foul</em></td>
<td>ūi</td>
<td>ūi</td>
<td>ūi</td>
<td>ūi</td>
</tr>
</tbody>
</table>
The Great Vowel Shift also explains what might seem like senseless allomorphy, such as that of the noun roots in *moon/mon-thly* and *goose/gos-ling*. The vowels in *moon* and *goose* were long, as we can still see from their spelling. During the Great Vowel Shift, they rose from [o:] to [u:]. The vowels in *month* and *gosling*, on the other hand, had been shortened, because they occurred before consonant clusters. They did not participate in the shift.

Changes in sounds and sound systems are constantly occurring in languages all over the world (see Sidebar 12.7). We cannot predict exactly which changes will occur at a particular time, but we do know about the kinds of changes that tend to occur under particular circumstances. Some of the more common types of sound change are listed in Textbox 12.1.

**TEXTBOX 12.1 SOME COMMON SOUND CHANGES**

Certain types of sound changes occur more often than others, usually because they result in less effort for speakers. Some of the more common types are below.

1. **Lenition**

   Speakers often try to put as little energy into pronunciation as possible. Some common kinds of lenition, or weakening, are shifts of voiceless stops to voiced ones (p > b, t > d, k > g), stops to fricatives (p > f, t > s, k > x), oral fricatives to h (f > h, s > h, ð > h, x > h), voiced obstruents to glides (b > w, v > w, g > j, dz > j), and shifts of vowels toward the center (i > ǝ, e > ǝ, a > ǝ).

   We saw an example of lenition in the shift of Old English *dæg* to Modern English *day*.

2. **Loss**

   Extreme weakening ultimately results in complete loss of a sound. We saw loss of the velar stop [k] from English words like *knee* and *knife*, and loss of the velar fricative [x] from English words like *night* and *right*. The lenition and loss of final [e] is rampant, as in *her* and *there*.

3. **Consonant addition: excrescence**

   Consonants are sometimes added sporadically. Proto-Indo-European *swer-* developed into Proto-Germanic *swestr*, with an added t, the ancestor of Modern English *sister*. Old English *junnrian* changed into *jundrian*, ‘thunder,’ with an added d. (Compare German *Donner*.)

   Such excrescent consonants usually result from tiny shifts in the timing of articulatory movements. Moving from the nasal stop /n/ to the oral continuant /r/ of *junnrian*, speakers closed off the nasal passage before opening up the mouth, resulting in the oral stop /d/.

4. **Vowel addition: prothesis and epenthesis**

   Vowels are sometimes inserted to break up consonant clusters. Addition at the beginning of a word is termed prothesis. A well-known example is the shift of Latin *spiritus* to Spanish *espiritu* ‘spirit.’ Addition within a word is termed epenthesis, as in the pronunciation of English *athlete* as *atha*lete.

5. **Fusion**

   Two sounds sometimes merge into one. We saw an example of vowel fusion in the shift of *Old English died* to modern *death*, now pronounced *[dɛθ]*.

6. **Breaking**

   One vowel sound may separate into two. We saw breaking in the long high vowels affected by the Great Vowel Shift. An example is the vowel of English *foul*, pronounced *[uː]* in Chaucer’s time but *[au]* today.
12.3 Grammar

It might be tempting to think that since all languages have grammar, often with fairly rigid rules, the grammar we see now has been there from the beginning. But like sounds, grammar is constantly evolving.

12.3.1 Grammaticalization processes

Grammatical meanings are typically expressed in languages with small words (e.g., articles, prepositions, auxiliaries, etc.), clitics (e.g., English genitive 's, future 'll, and negative n't), and affixes (prefixes, suffixes, infixes, circumfixes). Some examples of affixes discussed in other chapters are given in (2).

(2) Some grammatical affixes

<table>
<thead>
<tr>
<th>Language</th>
<th>Affixes</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manange</td>
<td>a-kʰa₂²²</td>
<td>'not come'</td>
</tr>
<tr>
<td>Goemai</td>
<td>mòe-japnuún</td>
<td>'siblings'</td>
</tr>
<tr>
<td>Karo</td>
<td>o-yàwan</td>
<td>'I left'</td>
</tr>
<tr>
<td>Seneca</td>
<td>sa-ayö</td>
<td>'he came back'</td>
</tr>
<tr>
<td>Chontal</td>
<td>sago-duy</td>
<td>'eating'</td>
</tr>
<tr>
<td>Ilocano</td>
<td>gilin-an</td>
<td>'grinder'</td>
</tr>
</tbody>
</table>

The grammatical morphemes that arise in languages are not random. They grow out of what speakers choose to say most often on an everyday basis. Most grammatical morphemes begin life as full words, usually with relatively general meanings. In some cases, these origins can still be detected. The English grammatical morphemes beside and besides, for example, developed from Old English phrases be sidan and be sides 'by the side' (OED 818). The word be is the ancestor of our modern preposition by. The word sidan is the Old English noun side,
with the dative case ending (required by the preposition be). The word sides is the same noun with the genitive case. Both originally only had a spatial meaning: ‘by the side of, near.’ The first can still be used that way: beside the tree. The second is now used only with more a abstract meaning: ‘in addition, moreover.’ The development of lexical forms to grammatical forms (e.g., prepositional phrases to prepositions and adverbs), and from grammatical to even more grammatical forms (e.g., the abstract adverb besides) is termed grammaticalization.

Grammaticalization does not affect single words in isolation. It normally applies to words in particular constructions. As speakers use a certain sequence of words with increasing frequency, it can become automated and processed as a unit. It is a bit like driving to a friend’s house in an unfamiliar area. The first time you go, you might think consciously about each stage of the journey: turning off the highway here, continuing for one mile, veering left, passing the fire station, etc. After several visits, you no longer focus on the individual steps; you simply drive to visit your friend. A similar cognitive process occurs with frequent phrases or constructions. Rather than focusing on each word; speakers simply select the whole construction as a chunk. Grammar is the product of the cognitive routinization of frequently recurring patterns of expression.

Constructions undergoing grammaticalization are typically extended to more and more contexts. As a result, their meanings tend to become more general and more abstract. We saw an example of such abstraction in the shift of the meaning of besides from physical location to ‘moreover.’ (See Sidebar 12.8 for another such example.) In most languages, full lexical items such as nouns and verbs occur with certain inflectional morphemes, such as tense or plural markers. These markers typically disappear during grammaticalization, a process termed decategorialization. We can see this in the development of English auxiliaries, which no longer carry person agreement. We can say She sing-s, but not She can-s sing. Finally, words undergoing grammaticalization within a particular construction tend to lose their individual salience, typically losing their stress and ultimately showing further phonological erosion. This, too, we can see in the case of English auxiliaries, such as the erosion of will to ‘ll, as in I’ll.

Prefixes and suffixes usually develop through similar processes. We can still perceive the origin of some. The English suffix -ful in playful, masterful, and useful creates adjectives. In measure terms like spoonful, cupful, and armful, it creates nouns. The source of the suffix is still obvious: it developed from the adjective full, which survives in present-day English. When a word undergoes grammaticalization, the original form may continue in the language with its function. Both the adjective full
and the suffix -ful already existed in Old English. The two have survived alongside each other for over a thousand years.

We can observe processes of grammaticalization occurring in English as we speak. One of the most discussed examples is /aimns/ (see, e.g., Hopper and Traugott 2004:2–4). Do you recognize it? Try pronouncing it with a verb /aimns it/. This is I am going to, as in I am going to eat. This construction apparently developed out of constructions like I am going to see my father. The original meaning involved a change in physical location for a purpose. A locative adverb could be included: I am going [to town] to see my father. The intended action would occur in the future, following the change in location indicated by the motion verb go. The current meaning of the construction is simply future tense. This shift in meaning from purpose to futurity is a kind of metonymic process, or a change resulting from the frequent association of two elements in the same speech situation. (Purposes are associated with the future because of the future actions intended to bring them about.) The original construction normally occurred in contexts involving both purpose and futurity. Futurity was then reinterpreted as its primary meaning.

Grammaticalization can also have a syntactic side. At a certain point, speakers apparently reinterpreted the syntactic structure of the construction be going to from a main clause I am going plus a purposive subordinate clause to see my father (I am going [to see my father]), to a sentence with a future auxiliary phrase be going to: I am going to [see my father]. This process is called syntactic reanalysis. Speakers came to conceive of the syntactic structure in a new way. Such a shift is not immediately obvious. It becomes apparent only when these speakers create new sentences that are incompatible with the old analysis. When one says I am going to like it, it is unlikely that motion is intended. The fact that syntactic reanalysis has occurred is confirmed by sentences with a second verb go, like I am going to [go to town].

The be going to future construction also shows phonological effects that are typical of developing grammatical constructions: the words are losing their individual stress and undergoing further phonological erosion. We now hear I'm going to eat, I'm gonna eat, Imma eat, and even sometimes Inna eat. Note that such reduction does not occur with the original more concrete construction. No one would reduce I'm going to town to I'm wanna town.

Grammaticalization can involve metaphorical extension, or the use of an existing word or construction from one domain to express a concept in another. A frequent metaphorical change involves the extension of markers denoting spatial relations to use for the realm of time. An example can be seen in Cherokee, an Iroquoian language now spoken primarily in North Carolina and Oklahoma. Cherokee contains an ancient verb prefix ta- meaning 'hither, toward the speaker.'

(3) Cherokee cislocative ‘hither’ (Montgomery-Anderson 2008: 328, 354)

- ta-kinatansinoheéli ‘He’s crawling toward us’
- ta-ákiiluhcheéli ‘He came up to us’
This prefix has now been extended to indicate future tense as well.


\[
\begin{align*}
ta\text{-yuúhali} & \quad \text{‘He will look for it’} \\
ta\text{-kintlecheéli} & \quad \text{‘He will take revenge on us’}
\end{align*}
\]

This development suggests a view of the future as something that is coming toward us.

As we look at the kinds of distinctions encoded in grammatical markers in languages around the world, we find that some meanings occur in language after language, such as negation, tense, plurality, and causation. **The frequency of such grammatical morphemes points to certain universal human concerns, and to concepts that people tend to express often.** At the same time, we sometimes find surprising grammatical markers, with very specific meanings. The suffixes in (5) are from Nuuchahnulth.

(5) Nuuchahnulth suffixes (Stonham 2005)

\[
\begin{align*}
\text{‘ahs} & \quad \text{‘in a vessel or container’} \\
[tuw\text{‘-ahs-ii}] & \quad \text{‘jumped into the canoe’} \\
[tuw\text{‘-}] & \quad \text{‘jump’} \\
\text{‘aci-} & \quad \text{‘in the bay, gulf, inlet’} \\
[jii\text{‘-aci-ii}] & \quad \text{‘he was shooting in the bay’} \\
[jii\text{‘-}] & \quad \text{‘shoot’} \\
\text{‘at} & \quad \text{‘out of the woods’} \\
[wika\text{‘-at-’as}] & \quad \text{‘she did not come out of the woods’} \\
wika\text{‘-} & \quad \text{‘come’} \\
\text{‘ač’r’ul} & \quad \text{‘from snout to dorsal fin’} \\
hilwee\text{‘in su’č’iit-‘ač’r’ul} & \quad \text{‘it was five fathoms from snout to dorsal fin’} \\
[su’č\text{‘-}] & \quad \text{‘five-fathom’}
\end{align*}
\]

Nuuchahnulth is a language of the Wakashan family, spoken on Vancouver Island, off the coast of British Columbia. The ocean has been a central part of the lives of Nuuchahnulth people for a long time. Examples like these show that the general processes by which grammatical morphemes and patterns develop are very similar across languages, but the specific distinctions they encode are shaped by the concerns of individual speech communities.

12.3.2 Renewal: restoring expressive power

The kinds of processes seen in the previous section are constantly at work in all languages. But if this is the case, shouldn't all languages have eroded to nothing or nearly nothing by now? In fact, there are other processes of change that help to maintain a certain equilibrium.
Language change: linguistic dynamism

As frequently used expressions become routine, they can lose their expressive punch: their impact can fade. **But a primary function of language is communication: speakers use language to convey their thoughts, often in creative ways.** This creativity can restore freshness and power to the language. An illustration of such a cycle is the development of negative constructions in English. In Old English, negation was usually indicated with a negative particle *ic ne ic ne wat* ‘I don’t know.’ (The verb *wat* ‘know’ has since fallen out of English.) Negation is expressed frequently in all languages, so negative constructions are likely candidates for routinization and erosion over time. Old English *ne* was already a small word. But negation is crucial information. To highlight its importance, speakers often reinforced negative sentences with extra words, such as *wiht* ‘something, anything’ or *na:wiht* ‘nothing, not anything.’ Over time, the original negative marker *ne* eroded until it disappeared entirely. As a result, *na:wiht* was left as the only negative marker. With regular use, its emphatic force began to diminish as well. Its phonological form has also eroded to modern *not* and even *n’t.*

(6) English negation (Hock and Joseph 1996: 176)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Form</th>
<th>Meaning</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td><em>ic ne wat</em></td>
<td>‘I don’t know’</td>
<td>(Old English) basic</td>
</tr>
<tr>
<td></td>
<td><em>ic ne wat</em> (<em>na:wiht</em>)</td>
<td>‘I don’t know (no) thing’</td>
<td>emphatic</td>
</tr>
<tr>
<td>II</td>
<td><em>ic ne wat</em></td>
<td>‘I don’t know’</td>
<td>basic</td>
</tr>
<tr>
<td></td>
<td><em>ic ne wat</em> (<em>na:wiht</em>)</td>
<td>‘I don’t know at all’</td>
<td>emphatic</td>
</tr>
<tr>
<td>III</td>
<td><em>ic ne wat</em> (<em>na:wiht</em>)</td>
<td>‘I don’t know’</td>
<td>basic</td>
</tr>
<tr>
<td>IV</td>
<td><em>I wot(e) not</em></td>
<td>‘I don’t know’</td>
<td>(Shakespeare) basic</td>
</tr>
</tbody>
</table>

The cycle of **renewal** continues today. Modern speakers often reinforce what is now the ordinary negative *not* with words phrases like *at all* or *a bit.*

12.3.3 Analogy: repairing patterns

The human capacity for pattern recognition is crucial for learning and using a language. Not only do humans search for patterns as they learn their mother tongue, and use them as models for creating new utterances; they often go further to repair what they perceive to be irregularities. Language can change both when children make novel hypotheses about the patterns behind the speech they hear, and when speakers attempt to regularize existing patterns.

We saw earlier that sound change can create irregularities in morphology. The most common way to form plurals in English is simply to add a plural suffix to the basic form of the noun: *tree*/*trees.* But because of various events in the history of English, some nouns change their form in the plural, as in *leaf*/*leaves* / *li:f*/*li:v*. Stop here for a moment and listen to how you pronounce the pair *house*/*houses.* For many English speakers, the noun *houses* shows the same kind of allomorphy as *leaf*/*leaves*: it is /haus/ *hauz*-/. The cause is the same. But for other speakers, in many cases younger ones, the irregularity has been repaired: they use the same form of the noun root in the singular
and the plural: /haus/haus/. Repairs like this are termed analogical remodeling. Speakers perceive a strong pattern among certain pairs of words (tree/trees, flower/trees, bird/birds, cloud/clouds) and, by analogy to those pairs, remodel ‘exceptions’ to make them fit the pattern.

Whether the remodeling occurs when children are first acquiring their language or later in life, once it has occurred, the result simply becomes part of the language. Have you ever thought about the past tense of the verb dive? The original form is dived, but for many speakers, it is now dove. This might at first seem surprising; the most common past tense marker in English is -ed. But English also contains robust sets of what are called strong verbs. These verbs form their past tenses with a vowel change, a pattern called ablaut. One such pattern can be seen in drive/drove and ride/rode. The past tense of dive was apparently remodeled by analogy to such verbs.

In this section we have seen only a sample of the ways languages can evolve. Many more occur, some quite commonly, others more rarely. Most examples here have been taken from English, but all of these processes occur in languages around the world.

12.4 Language relationships

Example (7) shows the numerals ‘one’ through ‘five’ in a variety of languages, with a note on the transcription given in Sidebar 12.9. Take a moment to compare them. Can you organize the languages into groups, based on the forms of their numerals? Are there any languages that stand out as not belonging to any group?

(7) Numerals in twenty-seven languages

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>un</td>
<td>deux</td>
<td>trois</td>
<td>quatre</td>
</tr>
<tr>
<td>b.</td>
<td>uno</td>
<td>dos</td>
<td>tres</td>
<td>cuatro</td>
</tr>
<tr>
<td>c.</td>
<td>uno</td>
<td>due</td>
<td>tre</td>
<td>quattro</td>
</tr>
<tr>
<td>d.</td>
<td>um</td>
<td>dois</td>
<td>três</td>
<td>quatro</td>
</tr>
<tr>
<td>e.</td>
<td>un</td>
<td>doi</td>
<td>trei</td>
<td>patru</td>
</tr>
<tr>
<td>f.</td>
<td>eins</td>
<td>zwei</td>
<td>drei</td>
<td>vier</td>
</tr>
<tr>
<td>g.</td>
<td>een</td>
<td>twee</td>
<td>drie</td>
<td>vier</td>
</tr>
<tr>
<td>h.</td>
<td>ein</td>
<td>tsvei</td>
<td>drei</td>
<td>fier</td>
</tr>
<tr>
<td>i.</td>
<td>en</td>
<td>två</td>
<td>tre</td>
<td>fyra</td>
</tr>
<tr>
<td>j.</td>
<td>en</td>
<td>to</td>
<td>tre</td>
<td>fire</td>
</tr>
<tr>
<td>k.</td>
<td>jeden</td>
<td>dwa</td>
<td>trzy</td>
<td>cztery</td>
</tr>
<tr>
<td>l.</td>
<td>adin</td>
<td>dva</td>
<td>tri</td>
<td>četíře</td>
</tr>
<tr>
<td>m.</td>
<td>jeden</td>
<td>dvě</td>
<td>tìi</td>
<td>čtyri</td>
</tr>
<tr>
<td>n.</td>
<td>eden</td>
<td>dva</td>
<td>tri</td>
<td>četirë</td>
</tr>
<tr>
<td>o.</td>
<td>e’na</td>
<td>di’co</td>
<td>tri’a</td>
<td>te’sera</td>
</tr>
<tr>
<td>p.</td>
<td>egy</td>
<td>kettõ</td>
<td>három</td>
<td>négy</td>
</tr>
</tbody>
</table>
What could explain the similarities between these numeral words across languages?

Words in different languages may resemble each other for several reasons. One is onomatopoeia, or imitation of sounds. Names of birds, for example, are sometimes coined from imitations of their calls, like *whippoorwill* or *chickadee*. Speakers of different languages might come up with similar imitations. This is an unlikely explanation for similarities among numerals. A second possibility is chance: only a certain number of sounds can be made with the human mouth, and it is not impossible that the same combination of sounds could have similar meanings in different languages by chance. There is a verb *čʰúw* in Central Pomo, a language indigenous to Northern California, which means ‘eat.’ It has no relation to English *chew*, however. A third is language contact: often speakers adopt words from one language into another. This is the source of much English vocabulary, such as *spaghetti* (Italian), *champagne* (French), and even *tea* (Chinese). But words for the lowest numerals are rarely borrowed. The fourth and perhaps most common reason for lexical similarities across languages is that the languages are descended from the same parent language. The similar words are a common inheritance from their ancestor, that is, cognates.

*Languages are always changing.* The change typically begins with variation, from speaker to speaker and situation to situation. As long as speakers are communicating with each other, they will not change their speech so radically as to interrupt intelligibility. *When a community splits, however, and the splinter groups no longer interact, their speech will no longer change in parallel.* The longer they are separated, the greater the differences will become. Dialects will develop: people may still understand each other but notice differences, as in the case of British and American English. In time, speakers in the different communities will no longer understand each other, as in the case of English and German. At that point, they are said to speak different languages. All languages that are descended from a common parent are said to belong to the same language family. The languages within a family are often called daughter languages. They are said to be genetically related.
The languages represented in Example (7) are the following: (a) French, (b) Spanish, (c) Italian, (d) Portuguese, (e) Romanian, (f) German, (g) Dutch, (h) Yiddish, (i) Swedish, (j) Danish, (k) Polish, (l) Russian, (m) Czech, (n) Macedonian, (o) Greek, (p) Hungarian, (q) Finnish, (r) Turkish, (s) Indonesian, (t) Ilocano, (u) Arabic, (v) Hebrew, (w) Japanese, (x) Swahili, (y) Mohawk, and (z) Seneca. Most (a–o) are from the same language family as English, called Indo-European. Several, however, are from different families, and their numerals look quite different. Hungarian and Finnish (p, q) are from the Finno-Ugric family. Turkish (r) is from the Turkic family. Indonesian and Ilocano (s, t) are from the Austronesian family. Arabic and Hebrew (u, v) are from the Semitic branch of Afroasiatic. Mohawk and Seneca (y, z) are Iroquoian. Turkish (r), Japanese (w), and Swahili (x) have no relatives on the list.

Among the Indo-European languages, numerals in some languages resemble each other especially closely, such as those in French, Spanish, Italian, Portuguese, and Romanian (a, b, c, d, e); those in German, Dutch, and Yiddish (f, g, h); those in Swedish and Danish (i, j); and those in Polish, Russian, Czech, and Macedonian (k, l, m, n). As far as is known, the original Indo-European speech community separated into over a dozen groups: Romance (a–e), Germanic (f–j), Slavic (k–n), etc. Many of these groups then divided again. The Germanic group split into a North Germanic subgroup (the modern Scandinavian languages), a West Germanic subgroup (English, Dutch, Frisian, German, Yiddish), and an East Germanic subgroup (Gothic). In general, languages that have split the most recently show the most similarities: they have shared a longer history of common development.

12.4.1 Family trees

Relationships are often illustrated with what is termed a family tree or Stammbaum (see Sidebar 12.10). A sample fragment of the Indo-European family tree is given in Figure 12.1.

Indo-European, the language from which English, German, French, Russian, Greek, Albanian, Armenian, Farsi, Hindi, and many other languages are descended, is thought to have been spoken around the fifth millennium BCE. Because there are no written records of it, all that is known is what can be reconstructed by comparing the daughter languages (see Textbox 12.2 for more on methods used to determine subgroupings). A reconstructed ancestral language is termed a proto-language. The reconstructed ancestor of English and its relatives is called Proto-Indo-European.
**TEXTBOX 12.2 CLADISTICS**

Computational methods are also being explored for answering questions about subgrouping, that is, interrelationships among languages already known to have developed from a common ancestor. One method, which comes from work in evolutionary biology, is cladistics. Similarities among languages are calculated in terms of the numbers of features (or “characters”) they share, such as lexical cognates, sound changes, and inflectional morphology. A group of languages that share a significant number of features is called a “clade.” A clade can be likened to a subgroup in a traditional family tree. But while family trees show a common parent language at the top, with successive splits into subgroups and ultimately individual languages at the bottom as in Figure 12.1, cladistics first produces networks, or “unrooted trees.” Cladograms simply arrange languages in terms of degrees of similarity. An example of a cladogram, from Taylor, Warnow, and Ringe (1998: 400), is below. It shows degrees of similarity among some Indo-European languages, based on a set of forty-six characters.

The lengths of the lines here are meaningful. Note that the line representing the Avestan-Vedic branch is longer than that for the Albanian-Old English branch. This is meant to indicate that Avestan and Vedic underwent more common changes than Albanian and Old English.

![Figure 12.2](image_url)
12.5  The comparative method

To determine that languages are related, we begin by looking for resemblances among basic words, like the numerals above. But *random similarities do not, on their own, constitute evidence of genetic relationship*. The similarities could be due to any of the four factors mentioned above: onomatopoeia, chance, contact, or common inheritance. To uncover genetic relationships, we look for *recurring sound correspondences*. This is done by applying the **comparative method**. The method will be illustrated here with three languages indigenous to northern California: Wintu, Nomlaki, and Patwin.

(8) Wintu, Nomlaki, and Patwin (Shepherd 2006)

<table>
<thead>
<tr>
<th>Wintu</th>
<th>Nomlaki</th>
<th>Patwin</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ‘ear’</td>
<td>mat</td>
<td>mat</td>
</tr>
<tr>
<td>b. ‘milkweed’</td>
<td>boq</td>
<td>boq</td>
</tr>
<tr>
<td>c. ‘frog, toad’</td>
<td>wataqmet</td>
<td>wataq</td>
</tr>
<tr>
<td>d. ‘house’</td>
<td>qewel</td>
<td>qewel</td>
</tr>
<tr>
<td>e. ‘bone’</td>
<td>paq</td>
<td>paq</td>
</tr>
<tr>
<td>f. ‘wild goose’</td>
<td>laq</td>
<td>laq</td>
</tr>
<tr>
<td>g. ‘play, gamble’</td>
<td>laqal</td>
<td>laqal-a</td>
</tr>
<tr>
<td>h. ‘hole’</td>
<td>holoq</td>
<td>holoq</td>
</tr>
<tr>
<td>i. ‘hawk’</td>
<td>qačit</td>
<td>qačit</td>
</tr>
<tr>
<td>j. ‘good, straight’</td>
<td>čal-a</td>
<td>čal-a</td>
</tr>
<tr>
<td>k. ‘red-tailed hawk’</td>
<td>lad</td>
<td>čeqeq</td>
</tr>
<tr>
<td>l. ‘hurt, sick’</td>
<td>koy-i</td>
<td>koy-a</td>
</tr>
<tr>
<td>m. ‘be ashamed’</td>
<td>kay-a:</td>
<td>kay-a</td>
</tr>
<tr>
<td>n. ‘enemy’</td>
<td>yuken</td>
<td>yuken</td>
</tr>
<tr>
<td>o. ‘body hair’</td>
<td>sekey</td>
<td>sekey</td>
</tr>
<tr>
<td>p. ‘climb’</td>
<td>dek-ma</td>
<td>dek-na</td>
</tr>
<tr>
<td>q. ‘woodpecker’</td>
<td>tarak</td>
<td>tarak</td>
</tr>
<tr>
<td>r. ‘basket root’</td>
<td>sek</td>
<td>sek</td>
</tr>
<tr>
<td>s. ‘ash tree’</td>
<td>irik-mi</td>
<td>irit</td>
</tr>
<tr>
<td>t. ‘strap, tumpline’</td>
<td>su:t</td>
<td>su:t</td>
</tr>
<tr>
<td>u. ‘belt’</td>
<td>tir-i</td>
<td>tiri</td>
</tr>
</tbody>
</table>

12.5.1  Step I: identifying similar morphemes

The first step is to identify potential cognates, or morphemes that are similar enough in form and meaning to be descended from the same morpheme in a common parent language. Compare the words in (8) to see whether any might not belong. Wintu *lad* (k), a kind of hawk, is completely different in form from Nomlaki *čeqeq* and Patwin *tektek*. We will not consider it further in our comparison.

Potential cognates must also be similar enough in meaning to have developed from the same word in the parent language. Full semantic details of the words could not be included in (8) for reasons of space; however, some of the words in (8) have exactly the same meanings, such as *mat/mat/mat* ‘ear’ (a), while others differ. Differences do not
mean they cannot be related: meanings can change over time. But the differences must represent plausible semantic change. Nomlaki *yuken* and Patwin *yuken* in (n) are translated ‘enemy.’ Wintu *yuken* is translated ‘dangerous, at war, Shasta.’ The Shasta people were the northern neighbors of the Wintu. The semantic relationship between ‘enemy’ and ‘dangerous/Shasta’ is sufficient for us to consider them possible cognates. On the other hand, Nomlaki has a word *leni* ‘grass,’ and Patwin a word *leni* ‘yesterday.’ This difference is too great for them to be cognates.

Finally, it is crucial that specific morphemes be compared. The ‘frog’ words in (c) are Wintu *wataqmet*, Nomlaki *wataq*, and Patwin *watâk*. The element *-met* in Wintu is an ‘amphibian reptile suffix’ (Pitkin 1985: 348). It also appears in Wintu *yir-met* ‘mountain lizard’ and *yohâl-met* ‘(another kind of) frog.’ We only need to compare the roots *wataq*/*watâk*. In the forms in (p) for ‘climb,’ *dek-ma*, *dek-na*, and *det-mu*, only the roots *dek*-are comparable. We do not need to consider the suffixes in our comparison.

Most of this first step has already been done for you in (8). Potential cognates have been assembled, and morphemes have been separated with hyphens.

12.5.2 Step II: listing correspondences

When we compare the words *ma:t*/*ma:t*/*ma:t* ‘ear’ in (a), we see that where Wintu has *m*, Nomlaki and Patwin also have *m*. We can summarize this in what is called a correspondence set: *m*/*m*/*m*. In the same word we see *a/*/a:/a: and *t/*/t:/t:. In *qačit/qacit/katit* ‘hawk’ (i), we find the sets *q/q/k*, *a/a/a*, *č/č/t*, *i/i/i*, and *t/t/t*. These sets recur in the data, and with more data, we would see even more. Take a moment now to list all the consonant correspondence sets you find in the data in (8). List each set once, arranging any that share sounds, or are phonetically similar, near each other.

In the words ‘tumpline’ (t) and ‘belt’ (u), Nomlaki shows nothing where Wintu and Patwin both show *r*. We write this correspondence using a zero: *r/*/Ø/*r*. Nomlaki words for ‘woodpecker’ (q) and ‘ash’ (s) are not available. We represent this gap in the data with a hyphen when first setting up the sets. For the first sound in ‘woodpecker,’ we write *t/-/t*. Since we have only one other set where this pattern would appear to fit, in this case *t/t/t*, we can assume that *t/-/t* exemplifies this pattern. (On the other hand, if we had two other correspondence sets where the pattern could fit, such as *t/t/A* and *t/d/t*, then we could not make such an assumption.)

| p/p/p | s/s/s | y/y/y | m/m/m | l/l/l | q/q/k |
| b/b/b | h/h/h | w/w/w | n/n/n | i/i/i | k/k/č | č/č/t |
| d/d/d | r/Ø/r | k/k/t | t/t/t |

12.5.3 Step III: first pass at reconstruction

We now propose sounds in the parent language, Proto-Wintun, that could be the ancestors of each correspondence set. Each reconstructed sound should be such that:

a. The changes from it to each of its reflexes (descendent sounds) in the daughter language are as plausible as possible, i.e., the kinds of natural changes we know that
sounds undergo cross-linguistically. Earlier in this chapter we saw some particularly common kinds of change, e.g., lenition and assimilation.

b. The changes are as few in number as possible.

The most likely ancestor of the set \( p/p/p \), for example, is \( *p \). This is a hypothesis that Proto-Wintun contained a sound \( *p \) that remained unchanged during the several thousand years of development into modern Wintu, Nomlaki, and Patwin. We follow the same procedure for other uniform sets: \( *b \) for \( b/b/b \), etc. Each has remained the same in the daughter languages.

Proto-Wintun

\[
\begin{align*}
* p & \rightarrow p \text{ in W, N, P} \\
* b & \rightarrow b \text{ in W, N, P} \\
* t & \rightarrow t \text{ in W, N, P} \\
* d & \rightarrow d \text{ in W, N, P} \\
* s & \rightarrow s \text{ in W, N, P} \\
* h & \rightarrow h \text{ in W, N, P} \\
* m & \rightarrow m \text{ in W, N, P} \\
* n & \rightarrow n \text{ in W, N, P} \\
* l & \rightarrow l \text{ in W, N, P} \\
* w & \rightarrow w \text{ in W, N, P} \\
* y & \rightarrow y \text{ in W, N, P}
\end{align*}
\]

The same procedure gives us easy reconstructions for the vowels, which are generally identical in the daughter languages. So far things look so simple that we hardly need to write them out.

For the set \( r/\emptyset/r \), a reconstruction of \( *r \) is most appropriate, but this entails a change. The data here indicate that \( *r \) has disappeared in Nomlaki.

\[
* r \rightarrow r \text{ in Wintu, Patwin} \\
\emptyset \text{ in Nomlaki}
\]

We next turn to \( q/q/k \). Two possible reconstructions spring to mind: \( *q \) or \( *k \). A choice of \( *q \) entails just one change, \( *q > k \) in Patwin. It is a plausible change: shifts from uvular to velar are common cross-linguistically.

\[
* q \rightarrow q \text{ in W, N} \\
* k \rightarrow k \text{ in P}
\]

A choice of \( *k \) for the \( q/q/k \) set, however, would entail two changes, \( *k > q \) in Wintu and \( *k > q \) in Nomlaki. Moreover, this change is much less common cross-linguistically. (It occurs mainly by assimilation to low back vowels.)

12.5.4 Step IV: combining sets

We now turn to the sets \( k/k/e \) and \( k/k/t \). A separate proto sound must be posited for each correspondence set, unless there is evidence for combining the sets. A likely ancestral sound for both \( k/k/e \) and \( k/k/t \) would be \( *k \). There may have been one original sound, \( *k \), which developed one way in some contexts and another way in others. \( *k \) may have developed into Patwin \( e \) in some contexts, and \( t \) in others. To determine whether this happened, we investigate the environments where each occurs. (This procedure is similar to that used for finding allophones in
Language change: linguistic dynamity

Complementary distribution. We can list their contexts as follows. The blank represents the correspondence set in each word. The sound that occurs before it in each language is on the left, in the order Wintu/Nomlaki/Patwin. The sound that occurs after it in each language is on the right.

<table>
<thead>
<tr>
<th>/k/</th>
<th>k/k/t</th>
<th>č/k/k/t</th>
</tr>
</thead>
<tbody>
<tr>
<td>a/a</td>
<td>a/a</td>
<td>e/e/e</td>
</tr>
<tr>
<td>#/#/</td>
<td>#/#/</td>
<td>o/o</td>
</tr>
<tr>
<td>#/#/</td>
<td>#/#/</td>
<td>a/a</td>
</tr>
<tr>
<td>u/u</td>
<td>u/u</td>
<td>e/e</td>
</tr>
<tr>
<td>e/e</td>
<td>e/e</td>
<td>e/e</td>
</tr>
</tbody>
</table>

We do have complementary distribution. The set /k/č always occurs before vowels. In contrast, the set /k/k/t never occurs before vowels; it always occurs before a consonant or at the end of the word. With this information, we can posit a single ancestral sound *k. This *k remained k in Wintu and Nomlaki. It changed to č (the affricate [tʃ]) in Patwin. This is a plausible sound change, a simple fronting of the obstruent. Before another consonant or at the end of a word, this č was simplified to the stop t. This too is a plausible change.

*č > t in Patwin

We are left with the set č/c/t. The most obvious reconstruction is *č, a sound not used for our Proto-Wintun consonant inventory so far. It implies just one change, *č > t in Patwin, and, as noted, it is a plausible change.

We should check to be certain that there is not complementary distribution between the contexts in which the č/c/t and the t/t/t correspondences occur. In fact there is not. Both occur, for example, at the beginning of a word before a, as in (i) ‘good, straight’ (č/c/t), and (q) ‘woodpecker’ (t/t/t). There is also no complementary distribution between the sets k/k/t and t/t/t. Both occur after e at the end of a word, as in (r) ‘root’ (k/k/t) and (a) ‘ear’ (t/t/t). A check of the contexts for the sets l/l and t/t/t also shows no complementary distribution. Both occur, for example, at the beginning of words before a, as in (f) ‘wild goose’ (l/l) and (g) ‘play, gamble’ (l/l).

12.5.5 Step V: ordering rules

Each of our rules represents a hypothesis about a sound change that took place in the language. But did they all occur at once? That is unlikely. If not, can we tell anything
about the order in which they occurred? Looking back at the last section, we see that we hypothesized the two changes below for Patwin.

\*q > k  
\*k > \*č

Our work indicates that the Proto-Wintun word for ‘hole’ was *holoq. If we assume that the two rules we see there operated in that order (1 before 2), we would have the following sequence of changes:

\*holoq > *holok > holoč

\*q > k   \*k > č

If, however, we hypothesize that they occurred in the opposite order (2 before 1), we would have the following sequence of changes:

\*holoq > *holok

\*k > č \*q > k

Change b would not affect the word ‘hole,’ because at that point it did not yet contain a k.

12.5.6 Step VI: Inventory check

We now consider the inventory of consonants that we have reconstructed for Proto-Wintun:

\*p \*t \*č \*k
\*b \*d
\*s \*h
\*m \*n
\*w \*l \*r \*y

This is a reasonable, balanced inventory. The absence of \*g is noteworthy, but languages with \[b\] and \[d\] but no \[g\] are actually not uncommon cross-linguistically.

12.5.7 Step VII: reconstructing words

We are now in a position to reconstruct full words. We will assume that the vowel inventory consists of \*i, \*e, \*a, \*o, and \*u, with no changes in the daughter languages. Some reconstructions are straightforward:

<table>
<thead>
<tr>
<th>Wintu</th>
<th>Nomlaki</th>
<th>Patwin</th>
<th>Proto-Wintun</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ‘ear’</td>
<td>ma:t</td>
<td>ma:t</td>
<td>ma:t</td>
</tr>
</tbody>
</table>
Others require undoing the sound changes we posited:

b. ‘milkweed’ boq boq bok *boq
i. ‘hawk’ qačit qačit katit *qačit
t. ‘tumpline’ surut súc surut *surut

Where the meanings of the cognates in the daughter languages are not identical, we must choose a proto meaning. The words in (f) with their translations are Wintu laq ‘wild goose,’ Nomlaki laq ‘brant, green-wing teal,’ and Patwin laklak ‘goose species.’

The most likely meaning of the Proto-Wintun term is ‘goose.’

Comparative reconstruction is one of the primary methods that linguists have used to determine language relatedness and to reconstruct protolanguages. A complementary method of internal reconstruction is discussed in Textbox 12.3.

12.6 Linguistic paleontology

Reconstruction of vocabulary in a proto-language can provide glimpses of the world of its speakers. If, for example, we can reconstruct a term for ‘snow,’ we can hypothesize that the speakers lived in an area with snow or within sight of snow. We have this term for Proto-Indo-European, reconstructed...
Marianne Mithun

as *neigwh-. Watkins (1981) provides an overview of Indo-European life as seen through reconstructed vocabulary.

The Indo-Europeans apparently lived in an area with wolves (*wl𝑝, *wlkwo-), bears (*tskɔ-), beavers (*bhịbhịru-), and mice (*m̥uːs-). There were salmon or trout (*laks-), eels (*angwhi-), cranes (*ger-), eagles (*cr-), thrushes (*trozdɔs-), and starlings (storos). They knew wasps (*wopsa-), hornets (*k̥ps-ro-), flies (*muv-), and bees (*bheis). They gathered honey (*melit-) and made mead (*medlu-), a honey-based alcoholic beverage. Watkins hypothesizes that the Indo-Europeans oriented themselves by facing east, because the root *deks- ‘right’ also meant ‘south.’

Kinship terms reveal an interesting pattern. There are numerous terms for relatives by marriage on the husband’s side, such as ‘husband’s father,’ ‘husband’s mother,’ ‘husband’s brother,’ ‘husband’s sister,’ ‘husband’s brother’s wife,’ and ‘son’s wife,’ but none for the corresponding terms on the wife’s side. This fact suggests that couples took up residency with the husband’s family.

The Indo-Europeans were farmers. They had words for *grain (‘grono, ancestor of our corn), or wheat or spelt, and perhaps rye and barley. They ground it (*melös, ancestor of meal and mihl). They had verbs for ‘sow’ (*sɛ-, ‘plow’ (*plɔɡ-), ‘yoke’ (*yeug-), and ‘gather’ (*kerp, ancestor of harvest). They kept livestock, including cattle (*gwoʊ, ‘cow/bull’), sheep (*sɔw), lambs (*agwimhɔ), goats (*ghaidɔ)-, and pigs (*porko- ancestor of farrow). They had dogs (*kwon-), and this term was apparently the basis for their word for ‘horse’ (*ekwo-). The root *peku- meant both ‘wealth’ and ‘cattle.’ They could weave (*webh-), sew (*syū-), and produce textiles (teks- ‘fabricate, weave’). They knew the wheel (*k_walo, based on the verb root *k_w(H)- ‘turn’).

Indo-European vocabulary was rich in words for ideas, abstractions, and relations. There were numerous terms for mental activity, among them *men- (source of English mind). There were also terms for ‘king’ (*reɡ-), a deity (*deiw-, and religious law (*leg- and *yewo, ancestor of Latin ās), for preaching, praising, and prophesying or singing (*sengwh-).

As we reconstruct Proto-Indo-European vocabulary and draw inferences about the lives of the Indo-Europeans, we cannot help but wonder just where these people lived. There have been many hypotheses, but the puzzle is not yet solved. One approach is to compare the reconstructed vocabulary with what we know about the natural environment and archaeological findings for that period, around the fourth millennium BCE, probably not earlier than 5000 BCE and not later than 2500 BCE. The Indo-Europeans apparently lived where there were wolves, bears, beavers, foxes, otters, hedgehogs, and mice; sparrows, quail, thrushes, cranes, vultures, blackbirds, crows, ravens, eagles, jays, pheasants, and storks; and turtles, frogs, and snakes. But unfortunately most of these animals and birds are ubiquitous through Europe and adjacent Asia, so they do not help us to pinpoint a specific Indo-European homeland.

The term for ‘birch’ is clearly reconstructible to *bherg’o. As pointed out by Mallory (1989: 161), it denotes the birch in Indic (bhunja-), Iranian (Ossetic bārz),
Germanic (birch), Baltic (Latvian beržs), and Slavic (Russian berëza). But the Latin cognate fraxinus means ‘ash,’ and there is no cognate in Greek. About half of the tree names reconstructed for Proto-Indo-European show a shift in meaning in Greek. These facts are taken to suggest that the Indo-Europeans originated elsewhere, then later moved into the area, applying original terms to the new trees they encountered.

One possible location of the homeland is the grasslands area north of the Black and Caspian Seas, known as the Pontic-Caspian steppes (Fortson 2004: 41). The reasoning behind this proposal is interesting. We know that the Indo-Europeans had the wheel. Archaeological evidence suggests that wheeled vehicles were invented around 3300–3400 BCE. The Indo-European community could thus not have broken up before this time. At this time, the steppes were inhabited by a group known as the Yamna, who came from the area between the steppes and the nearby forest between the Dnieper and Volga rivers. We know that the Indo-Europeans knew horses. The teeth of horses found in this area from that time show microscopic abrasions, from clamping down on a bit. Additional aspects of culture reconstructed through linguistic paleontology match archaeological findings in the area. But without a written tradition, it is not possible to identify the language of communities uncovered archaeologically. This hypothesis, while promising, must remain just that.

**CHAPTER SUMMARY**

We have seen that all aspects of language undergo change. At the phonological level, certain types of sound change are particularly frequent, such as those involving lenition or assimilation. Many of these are motivated by a desire on the part of speakers to reduce the effort necessary for speech. Sound change can have effects beyond the shifting of individual sounds: it can result in the restructuring of the phonological system, with the addition and loss of distinctive features and phonemes. It can also produce seeming irregularities, such as allomorphy. Grammar tends to develop via certain pathways. Grammatical morphemes develop most often from independent lexical items, a process termed grammaticalization. Grammaticalization typically involves a constellation of changes, including cognitive routinization, generalization of meaning, decategorialization, and ultimately phonological erosion. But not all change is reductive. Speakers are constantly reanalyzing the patterns underlying speech, repairing perceived irregularities by analogical remodeling, and renewing constructions that have lost their expressive force.

The fact that languages are constantly changing can result in the differentiation of languages over time. A language family consists of all of those languages descended from a common parent or proto-language. Relationships among these languages are often represented by family tree diagrams. In this chapter we have seen an example of the comparative method, used to establish genetic relationships among languages and to reconstruct aspects of their common parent, such as its sound system and
vocabulary. The reconstruction of vocabulary in the proto-language can in turn provide glimpses into the lives of its speakers.

Languages are dynamic systems, constantly changing in all areas of their structure and content. The changes are not predictable, but they are also not random. They are shaped by certain human cognitive faculties such as pattern recognition, the routinization of recurring patterns of expression, and the creative acts of speakers eager to find fresh and powerful ways of conveying their thoughts.

**TEXTBOX 12.4 GLOSSING CONVENTIONS USED IN THIS CHAPTER**

<table>
<thead>
<tr>
<th>Convention</th>
<th>Meaning</th>
<th>Convention</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEG</td>
<td>negation</td>
<td>REP</td>
<td>repetitive</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
<td>DUR</td>
<td>durative</td>
</tr>
<tr>
<td>PST</td>
<td>past tense</td>
<td>NMLZ</td>
<td>nominalizer</td>
</tr>
</tbody>
</table>

**SUGGESTIONS FOR FURTHER READING**

**Slade, Benjamin** (ed.). Beowulf on Steorarume. www.heorot.dk/.

You can find the full text of *Beowulf* by searching for “beowulf” online. A variety of translations are available. You can learn more about the language of the time from an online glossary of all words in the manuscript.


This book is an excellent introduction to the history of English.


This book features a fascinating but somewhat advanced article on grammaticalization.


This book contains further discussion of the Proto-Europeans.

**EXERCISES**

In the exercises here, related forms have already been assembled for you. Normally a major part of work in historical linguistics involves this step as well, the discovery of related forms. These exercises provide just enough data for you to be able to detect certain patterns. Normally, one would consider much more data for a solid analysis.
1. French vowels: sound change  
(Thanks to Ben Fortson for expertise on the Romance exercises.)

For this problem, look only at the stressed syllables. First describe in words what happened to the vowels in these syllables during the development of French. Then write a rule to capture your description.


Languages of the Eskimo-Aleut family are spoken over a wide area of the Arctic, from Siberia to Greenland. The family consists of two main branches: Eskimoan and Aleut. The language of Greenland, called Kalaallisut, is a member of the first branch.

The data here show one consonant sound shift. Write a rule to describe this sound change.

3. Italian laterals: conditioned sound change

For this problem, look only at the development of the lateral / from Latin into Italian. First describe the change in words. Then write a rule that captures what happened.
4. French sibilants: conditioned sound change

The earliest documentation we have of French is from the ninth century. The term Old French designates stages of the language from that time up to the fourteenth century. Of course the language was undergoing change throughout that period.

The French sibilants underwent systematic changes during the thirteenth century. (Look at the IPA transcriptions of Old French and Modern French, rather than their standard spellings. Each of the sequences tʃ, ts, ʤ, and dz represents a single consonant sound, an affricate.)

i. Write rules to describe these changes. Describe in words the changes specified in each rule.

ii. Describe in words any other changes that you notice here.

iii. Can you understand why modern French spelling, in the right hand column, seems to fit modern French pronunciation so badly?

<table>
<thead>
<tr>
<th>Latinorthography</th>
<th>OldFrench</th>
<th>ModernFrench</th>
<th>Spelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>c, ch</td>
<td>tʃamptʃ</td>
<td>ją</td>
<td>champ</td>
</tr>
<tr>
<td>gh</td>
<td>tʃarʤtʃ</td>
<td>jare</td>
<td>charge</td>
</tr>
<tr>
<td>i after a consonant</td>
<td>ʤorsʤ</td>
<td>ʒorsʤ</td>
<td>Georges</td>
</tr>
<tr>
<td>k, ‘full’</td>
<td>plękļu</td>
<td>pino</td>
<td>pieno</td>
</tr>
<tr>
<td>l, ‘plant’</td>
<td>planta</td>
<td>pianta</td>
<td></td>
</tr>
<tr>
<td>m, ‘flat’</td>
<td>plńu</td>
<td>piano</td>
<td></td>
</tr>
<tr>
<td>n, ‘lead’</td>
<td>plumbu</td>
<td>piombo</td>
<td></td>
</tr>
<tr>
<td>p. ‘key’</td>
<td>clävis</td>
<td>chiave</td>
<td></td>
</tr>
<tr>
<td>q. ‘to close’</td>
<td>clüdĕre</td>
<td>chiudere</td>
<td></td>
</tr>
<tr>
<td>r. ‘flame’</td>
<td>flämma</td>
<td>fiamma</td>
<td></td>
</tr>
<tr>
<td>s. ‘flower’</td>
<td>fiøre</td>
<td>fiore</td>
<td></td>
</tr>
<tr>
<td>t. ‘river’</td>
<td>flämēn</td>
<td>fiume</td>
<td></td>
</tr>
<tr>
<td>u. ‘white’</td>
<td>blāncu</td>
<td>bianco</td>
<td></td>
</tr>
<tr>
<td>v. ‘glacier’</td>
<td>gläciāriu</td>
<td>giacciaio</td>
<td></td>
</tr>
<tr>
<td>w. ‘double’</td>
<td>dupļu</td>
<td>doppio</td>
<td></td>
</tr>
<tr>
<td>x. ‘example’</td>
<td>exemplu</td>
<td>esempio</td>
<td></td>
</tr>
<tr>
<td>a. ‘field’</td>
<td>tʃamp</td>
<td>ją</td>
<td>champ</td>
</tr>
<tr>
<td>b. ‘cow’</td>
<td>vatjo</td>
<td>vaj</td>
<td>vache</td>
</tr>
<tr>
<td>c. ‘judge’</td>
<td>dɛ̃yɾdʒ</td>
<td>jʒe</td>
<td>jage</td>
</tr>
<tr>
<td>d. (name)</td>
<td>dʒordʒas</td>
<td>sɔɾʒ</td>
<td>Georges</td>
</tr>
<tr>
<td>e. ‘charge’</td>
<td>tʃarʤer</td>
<td>jarge</td>
<td>charger</td>
</tr>
<tr>
<td>f. ‘hundred’</td>
<td>tsɔnt</td>
<td>są</td>
<td>cent</td>
</tr>
<tr>
<td>g. ‘hunt’</td>
<td>tʃatsʃer</td>
<td>jase</td>
<td>chasser</td>
</tr>
<tr>
<td>h. ‘shield’</td>
<td>eskɔy</td>
<td>eky</td>
<td>écu</td>
</tr>
<tr>
<td>i. ‘sword’</td>
<td>ɛspoɾdɔ</td>
<td>epe</td>
<td>épée</td>
</tr>
<tr>
<td>j. ‘our’</td>
<td>nastrɔ</td>
<td>notr</td>
<td>notre</td>
</tr>
<tr>
<td>l. ‘islands’</td>
<td>izlɔs</td>
<td>il</td>
<td>iles</td>
</tr>
<tr>
<td>k. ‘taken’</td>
<td>pris</td>
<td>pri</td>
<td>pris</td>
</tr>
<tr>
<td>l. ‘have’</td>
<td>avets</td>
<td>ave</td>
<td>avez</td>
</tr>
<tr>
<td>m. ‘to the’</td>
<td>alts</td>
<td>o</td>
<td>aux</td>
</tr>
<tr>
<td>n. (name)</td>
<td>frantiagoʃ</td>
<td>fraŋswa</td>
<td>François</td>
</tr>
</tbody>
</table>
5. Uto-Aztecan: comparative method (data from Sapir 1930; Lamb 1958; Miller 1972; Munro and Mace 1995; assembled in Miller 1988.)

The languages here are all from the Uto-Aztecan family. Shoshone territory stretches over areas of Nevada, Utah, Idaho, and Wyoming; Southern Paiute territory from southeastern California into Nevada, Arizona, and Utah; Mono territory on both sides of the Sierra Nevada Mountains in east central California; and Tubatulabal territory along the Kern River near modern Bakersfield, California.

A number of developments have occurred since these languages diverged. Some of the words here include morphemes that are not cognate. These are set off by hyphens. You do not need to include those non-cognate morphemes in your analysis.

For this problem, investigate only the velar stops and any related sounds. Reconstruct a proto-sound or sounds and write rules to describe developments in the various languages:

<table>
<thead>
<tr>
<th>Shoshone</th>
<th>Southern Paiute</th>
<th>Mono</th>
<th>Tubatulabal</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ‘fire, heat’</td>
<td>ku-</td>
<td>ku-</td>
<td>ku-na</td>
</tr>
<tr>
<td>b. ‘husband’</td>
<td>kuhma</td>
<td>kumma</td>
<td>kuwa</td>
</tr>
<tr>
<td>c. ‘neck’</td>
<td>kuta</td>
<td>kuta</td>
<td>kutta</td>
</tr>
<tr>
<td>d. ‘firewood’</td>
<td>kuna</td>
<td>kunna</td>
<td>kun(n)a</td>
</tr>
<tr>
<td>e. ‘no’</td>
<td>ke</td>
<td>ka</td>
<td>qa-tuʔu</td>
</tr>
<tr>
<td>f. ‘rat’</td>
<td>kaŋ</td>
<td>ka:</td>
<td>—</td>
</tr>
<tr>
<td>g. ‘jackrabbit’</td>
<td>kamma</td>
<td>kammi</td>
<td>qammi</td>
</tr>
<tr>
<td>h. ‘house’</td>
<td>kahni</td>
<td>kanni</td>
<td>—</td>
</tr>
<tr>
<td>i. ‘wing, feather’</td>
<td>kasa</td>
<td>kasa-p</td>
<td>qassa</td>
</tr>
<tr>
<td>j. ‘bite’</td>
<td>ki-</td>
<td>kiʔi-</td>
<td>ki-</td>
</tr>
<tr>
<td>k. ‘edge’</td>
<td>kiŋ-a:</td>
<td>kiwa</td>
<td>—</td>
</tr>
<tr>
<td>l. ‘pat grandfather’</td>
<td>kinu</td>
<td>kunn-</td>
<td>kinu</td>
</tr>
<tr>
<td>m. ‘come’</td>
<td>kimma</td>
<td>kimma</td>
<td>kimat</td>
</tr>
<tr>
<td>n. ‘elbow’</td>
<td>ki-</td>
<td>ki-ppi</td>
<td>ma-kipi</td>
</tr>
<tr>
<td>o. ‘break, cut in pieces’</td>
<td>—</td>
<td>kapi-</td>
<td>to-qopi</td>
</tr>
<tr>
<td>p. ‘face’</td>
<td>kopa-i</td>
<td>kopa-</td>
<td>qope</td>
</tr>
<tr>
<td>q. ‘snake’</td>
<td>tokoa</td>
<td>tokoa</td>
<td>toqoqq’a</td>
</tr>
<tr>
<td>r. ‘face’</td>
<td>kopa-i</td>
<td>kopa-</td>
<td>qope</td>
</tr>
</tbody>
</table>


New Guinea is home to a vast number of languages, estimated at over 850. They show tremendous diversity as well. Many are not yet described, or are under-described. The Gum languages are spoken in northeastern Papua New Guinea, in Madang Province.

For this problem, you will reconstruct just some of the consonants.

i. Make a list of all correspondence in the data that involve s.

ii. Make a list of all correspondence sets involving any of the other consonants in these s sets.

iii. Provide reconstructions for all of your correspondence sets.

(There is no complementary distribution here.)

iv. Using rules, list all sound changes that affected your proto consonants in each language.

v. Describe in words the rationale for each of your reconstructions.

<table>
<thead>
<tr>
<th>Gumalu</th>
<th>Amele</th>
<th>Bau</th>
<th>Panim</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ‘fingernail’</td>
<td>siu</td>
<td>hilo</td>
<td>si</td>
</tr>
<tr>
<td>b. ‘vein’</td>
<td>sirima-</td>
<td>hilima-</td>
<td>sirima-</td>
</tr>
<tr>
<td>c. ‘coconut tree’</td>
<td>asur</td>
<td>ahul</td>
<td>asur</td>
</tr>
</tbody>
</table>
7. Romance labials: comparative method

(There are additional complexities in the languages beyond those shown in the data here.)

Provide reconstructions for all of the labial sounds, and then write rules that specify the development of each into the modern language. You may need to compare the phonological environments in which similar correspondence sets occur.

SIDEBAR 12.15

TRANSCRIPTION NOTE

The material here is in the standard spelling systems for these languages.

<table>
<thead>
<tr>
<th>Romance orthography</th>
<th>IPA</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gn (Ital.), ñ (Spanish)</td>
<td>[ɲ]</td>
<td>palatal nasal</td>
</tr>
<tr>
<td>c (Spanish)</td>
<td>[k]</td>
<td></td>
</tr>
<tr>
<td>b (Spanish)</td>
<td>[β]</td>
<td></td>
</tr>
<tr>
<td>h in Spanish is unpronounced</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. Takanan Reconstruction (data from Girard 1971)

The Takanan languages are spoken in northern Bolivia and southeastern Peru. For this problem, you will do a complete reconstruction of the sounds of Proto-Takanan and some vocabulary.

i. Make a list of sound correspondences.
ii. Group the correspondence sets by phonetic features so that sets sharing sounds are adjacent.
iii. Check to see whether any related sets are in complementary distribution.
iv. Reconstruct a proto-sound for each set or each group of sets in complementary distribution.
   There should be a different proto-sound for each correspondence set or group of sets in complementary distribution. The proto-sound you choose for each should entail likely sound changes.
v. With each reconstructed sound, provide rules for its development in the modern languages.
vi. If any of the sound changes you list in your rules is a recognized kind of phonological process, name that process next to the rule.
vii. Check to see that you have reasonable consonant and vowel inventories.
viii. On the basis of your reconstructed sounds and rules, reconstruct the Proto-Takanan word that is the ancestor of each word above.
9. Grammaticalization

Each pair of sentences below shows evidence of certain grammaticalization processes within the language. Identify the forms that have undergone any of these processes, and name the processes involved in each development.

1. Bambara (data from Donald Lessau cited in Heine and Kuteva 2002:75) Bambara is a Mandé language spoken in Mali and Senegal, in West Africa.
   a. 3PL AUX dome
      ‘They come.’
   b. 3SG NEAR.FUT die
      ‘He will die (soon and/or surely).’

   (Tones are not represented here.)
   a. give ASP 1SG five CLF
      ‘He gave me five dollars.’
   b. write ASP one-CLF letter to 3SG.M
      ‘I wrote him a letter.’

   Negerhollands is a Dutch-based creole that was once spoken in the US Virgin Islands.
Marianne Mithun

a. ju lo: afo fa mi
   2SG go in.front of 1SG
   ‘You go in front of me.’

b. am a flig lo mi di flut
   3SG PRF fly away PREP DEF flute
   ‘He flew away with the flute.’