4 Morphology: what’s in a word?

KEY TERMS
- Morphology
- Morpheme
- Affix
- Compound
- Productivity
- Allomorph
- Lexeme
- Derivation
- Inflection
- Lexicalization

CHAPTER PREVIEW
We have seen how groups of sounds pattern in each language as sounds are combined to form words. Morphology is the study of the internal structure of words. This chapter describes the kinds of building blocks used to form words in different languages and the ways they can be combined. Speakers are not usually conscious of the patterns inside of their words, but for most languages, knowledge of such patterns is crucial to the ability to speak. Topics discussed here include the notion of the word, the forms morphemes can take, the kinds of meanings they can carry, how to identify and describe them, and what we can learn from the morphological patterns we find. In the process, it introduces core grammatical concepts and analytical skills that will be central to discussions of word classes, syntax, and other topics covered throughout this book.

LIST OF AIMS
After reading this chapter, students should be able to:
- discuss ways to identify words in different languages;
- identify basic morphemes in an unfamiliar language;
- identify simple allomorphs and write rules to describe their distribution;
- identify the meanings of basic morphemes;
- define the term “compound” and give examples of compounds;
• explain the difference between possible and actual words;
• define the term productivity;
• explain what is meant by the terms “lexeme” and “lexicalization.”

4.1 Introduction to morphology

As speakers, what do we know about our mother tongue? Much of our linguistic knowledge is unconscious. When we start to learn a new language, we often come to appreciate just how much speakers must know in order to talk, whether they are aware of it or not.

It seems obvious that knowing words is a critical part of knowing a language, but what does it mean to know a word? A recent newspaper advertisement claims *Unbelievable weight loss breakthroughs!* Most English speakers recognize each of the words in this phrase as part of the vocabulary of English. They may or may not realize that words can have meaningful parts of their own. The first word, *unbelievable*, has three parts. It is based on the verb root *believe*. With the addition of the ending *-able* we have another word *believable* ‘able to be believed.’ If we add *un-* to that word *believable*, we have yet another word, this one with the opposite meaning: *un-believable*-able ‘not able to be believed.’ Each of the meaningful parts of a word, like *believe*, *-able*, and *un-* is called a **morpheme**. The study of how morphemes are combined to form new words is called **morphology**.

*Morphology is one of the areas in which languages can differ the most from each other.* Compare the two sentences below. The first is from Engenni, a Kwa language spoken in Nigeria. The top line in the example shows the sentence as it was spoken; the second line shows the meaning of each word, and the third line shows a free translation of the whole.

**SIDEBAR 4.1**

TRANSCRIPTION NOTE

<table>
<thead>
<tr>
<th>Engenni symbol</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>vowel with grave accent (e.g., ù)</td>
<td>low tone</td>
</tr>
<tr>
<td>vowel without an accent</td>
<td>high tone</td>
</tr>
<tr>
<td>vowel with acute accent (e.g., Á)</td>
<td>upstepped (extra-high) tone</td>
</tr>
</tbody>
</table>

(1) Engenni (Thomas 1978)

Áta na wa omù.
‘Let’s go look for the house.’

The second example is from Mohawk, a language of the Iroquoian family spoken in Quebec, Ontario, and New York State. Here again the top line shows the sentence as spoken; note that the whole sentence is expressed in a single Mohawk word. The second line shows the parts (morphemes) of the word. The third identifies the meaning of each morpheme. The fourth provides a free translation of the word as a whole.
Morphology: what’s in a word?

(2) Mohawk
Tenonhsihísá:kkha.
te-ni-nonhs-ihsak-ha
you.and.I-two-house-seek-go.and
‘Let’s go look for the house.’

The Engenni and Mohawk sentences mean roughly the same thing, but the two languages differ considerably in the way they distribute information over their words. In the Engenni example, each word contains just one meaningful part, or one morpheme. In the Mohawk example, all of the same information is packaged into a single word.

What is a word?

How do we know that the Mohawk sentence in (2) contains just one word? The most obvious way to find out is to ask a speaker. Mohawk speakers have clear ideas about where one word stops and the next begins, whether or not they have ever read or written their language. All would readily agree that (2) is a single word. Let’s look at a longer sentence. Compare the Mohawk sentence in (3) with its English translation. The amount of information contained in each word is different in Mohawk and English: the Mohawk sentence contains seven words, and the English translation fourteen. But all Mohawk speakers would again divide the sentence into words in the same way.

(3) Mohawk: Konwatsitsaiénni Rita Phillips, speaker
Ónen ki̊’ iáken’
ónen ki’ iak-en’
now in.fact one-say-s
now in fact one says
‘Now then, they say,
karhákon niha:titakenontie’
karh-ak-on n-i-a-ha:ti-take-n-ontie’
it-tree-be.in-is there-thither-FACTUAL-they-all-run-to.there-along
place in the forest they started running thither into it
they raced off into the forest
wahonhiá:khá’ sewahió:wane’;
wa-h-on-ah-ak-ha’ se-w-ah-owane’e’
FACTUAL-they-all-fruit-pick-go.and one-it-fruit-big-is
they went to pick fruit apple
to go pick apples.’

SIDEAR 4.2
TRANSCRIPTION NOTE
Mohawk symbol IPA
on [n]
en [ɛ]
: (colon) [:] (vowel length)
’ (apostrophe) [ʔ]
‘ (acute accent) high or rising pitch
(grave accent) falling tone
In some languages, the boundaries between some words are not as clear. For example, how many words are there in the English term *ice cream* or *ice-cream*? All languages change over time, and many aspects of language can change, including the boundaries between words. When two or more words occur together very frequently in speech, especially when they come to represent a single idea like ‘ice cream,’ speakers may begin to process them as a single unit, rather than composing the expression anew, word by word, each time they refer to it. The merging of words into a single cognitive unit can occur gradually over time. It is in such cases of change in progress that speakers may be uncertain of word boundaries.

Speaker intuitions usually provide the best indication of boundaries between words, although in some cases standardized spelling conventions can interfere with intuitions.) But individual languages may provide additional kinds of evidence. In most languages, words can be pronounced alone, but individual morphemes often cannot. The Mohawk -rh- in ‘forest’ above would never be pronounced by itself; speakers would not even recognize it if it were. Similarly, if I asked you what English *t* means, you probably would not recognize it as the past tense suffix at the end of *looked*.

Speakers can pause between words: *I... don’t really want to*. But they rarely if ever pause between the morphemes inside of a word: *Lightning was flashing in all directions*. This is true for Mohawk as well as for English. If Mohawk speakers are interrupted in the middle of a word, they go back and start over at the beginning. Speakers sometimes insert parenthetical information between words: *I... if you want to know the truth... don’t really want to*. But they do not insert such asides between morphemes: *Lightning... really scary... ing was flashing*. This suggests that *words are not usually produced morpheme by morpheme as speakers talk; words are stored as single cognitive units, and selected from memory as wholes for speech.*

In many languages, it is easy to identify words on the basis of phonological evidence. Perhaps the most obvious type of phonological evidence is the location of stress. Recall that the term “stress” refers to the relative prominence of a syllable in the word. A syllable with primary stress will often have special pitch, may be louder, and/or perhaps longer in duration than an unstressed syllable. In some languages, primary stress regularly falls on a certain syllable. In Finnish, for example, stress occurs on the first syllable of each word: *ääsi* ‘donkey,’ *perhonen* ‘butterfly,’ *välimatka* ‘distance,’ *valokuvauskone* ‘camera.’ In Spanish, primary stress (apart from specific exceptions) falls on the second-to-last syllable: *burro* ‘donkey,’ *mariposa* ‘butterfly,’ *biblioteca* ‘library.’ Mohawk shows the same pattern as Spanish. The accent marks in the Mohawk examples in (2) and (3) above indicate stressed syllables. *Other types of phonological evidence can provide clues to the boundaries between words as well, though these differ from language to language.*

Does wordhood matter? It can. Often what can be said in one word in a language can also be said, more or less, with multiple words in the same language. But there are usually differences in meaning, although they may be subtle. A number of the
morphemes in the Mohawk word in (4a) could be expressed with separate words. If, for example, you and I had been looking for something else beforehand, and I wanted us to shift the goal of our search, it would be more appropriate to name the house separately, as in (4b).

(4) Mohawk
a. *Tenínonsíhsákha.*  ‘Let’s go look for the house’
   let’s go house seek
b. *Kanónhsote’*  teníhsákha.  ‘Let’s go look for the house’
   it house stands let’s seek

Another example can be seen in Lakota, a Siouan language of the Plains. Both (5a) and (5b) could be translated ‘I made it black.’

(5) Lakota (Stan Redbird, speaker p.c.)
a. *Sábwyé.*  ‘I made it/them black.’
   I blackened
b. *Sápa wakágé.*  ‘I made it/them black.’
   black I made

When asked about the difference, the speaker, Mr. Redbird, explained that if he had just polished his boots he would use (5a). If he had hung a kettle over the fire to start a stew cooking, then returned some time later to find that the flames had come up and eventually left soot on the bottom of the kettle, he would use (5b). The first sentence indicates a single, direct action, while the second is appropriate for a longer string of events and indirect consequences.

**4.2 Kinds of morphemes**

Morphemes, the building blocks of words, can be classified in several ways. One distinction is between what are called free morphemes and bound morphemes. **Free morphemes are those that can stand alone as words.** In the English phrase we saw at the outset, the morphemes believe, weigh, loss, break, and through are all free, because they can be used as words on their own. All of the morphemes in the Engenni sentence in (1) are free: á ‘one,’ ta ‘go,’ na ‘to,’ wa ‘seek,’ and omà ‘house.’ In the Mohawk sentence in (3), the words ónen ‘now’ and ki ‘in fact’ are free. **Bound morphemes are morphemes that never occur as words on their own.** The English morphemes
un-, -able, and -s are all bound. Sidebar 4.4 explains the use of hyphens in writing out morphemes.

Another way to classify morphemes is into roots and affixes. Roots are considered the foundation of the word. Every word has a root at its core. The root usually conveys the main meaning of the word. The root of un-believ-able, for example, is believe.

Affixes are morphemes that attach to roots and modify their meaning in some way. The morphemes un-, -able, and -s are all affixes. Affixes never stand alone as words; they are bound (see Sidebar 4.5).

Affixes can also be classified according to where they occur: before the root, after the root, on both sides of the root, or inside the root. We can see each of these kinds of affixes in Ilocano, a language of the Austronesian family spoken in the Philippines.

A prefix is a type of affix that appears before the root, like the English un- of un-believable. Compare the Ilocano words in the left column in (6) with those on the right.

(6) Ilocano prefix (Rubino 1997)

- amianan ‘north’
- abagatan ‘south’
- bantay ‘mountain’
- ili ‘town’
- Amerika ‘America’

The prefix taga- forms words for origin or nationality from words for places.

A suffix is a type of affix that occurs after the root, like the English -able of un-believ-able. Compare the Ilocano words in the left column in (7) with those on the right.

(7) Ilocano suffix (Rubino 1997)

- giling ‘grind’
- sagat ‘strain’
- balkot ‘wrap’
- timbeng ‘weigh’

Adding the suffix -an to a verb can create a noun. Affixes that create nouns are called nominalizers. The nouns in (7) designate instruments. This particular suffix can thus be called an instrumental nominalizer.

An infix is a type of affix that appears inside of the root.
Ilocano infix (Rubino 1997)

(8) Ilocano infix (Rubino 1997)

\[
\begin{align*}
\text{kuton} & \quad \text{‘ant’} & \quad k\text{-}i\text{n}-u\text{ton} & \quad \text{‘ant-infested’} \\
\text{gayaman} & \quad \text{‘centipede’} & \quad g\text{-}i\text{n}\text{-}a\text{yaman} & \quad \text{‘infested with centipedes’} \\
\text{ngilaw} & \quad \text{‘fly’} & \quad n\text{g}-i\text{n}\text{-}i\text{l}a\text{w} & \quad \text{‘fly-infested’} \\
\text{kuto} & \quad \text{‘lice’} & \quad k\text{-}i\text{uto} & \quad \text{‘lice-infested’}
\end{align*}
\]

The Ilocano infix -in adds the meaning ‘infested.’ **When we describe an infix, we always specify where in the root it is inserted.** Here the infix is inserted after the first consonant of the noun: \(k\text{-}i\text{n}-u\text{ton}\) ‘ant-infested.’ (The sequence of letters \(ng\) is used in the practical spelling system to represent a velar nasal \([ŋ]\), a single consonant.)

Infixes are much rarer in the languages of the world than prefixes and suffixes.

Circumfixes wrap around the root, with part occurring before and part after.

(9) Ilocano circumfix (Rubino 1997)

\[
\begin{align*}
\text{rag}\text{sak} & \quad \text{‘happy’} & \quad \text{pag}\text{-}\text{rag}\text{sak}\text{-}e\text{n} & \quad \text{‘make someone happy’} \\
\text{leddaang} & \quad \text{‘sad’} & \quad \text{pag}\text{-}\text{leddaang}\text{-}e\text{n} & \quad \text{‘make someone sad’} \\
\text{uray} & \quad \text{‘wait’} & \quad \text{pag}\text{-}\text{uray}\text{-}e\text{n} & \quad \text{‘make someone wait’} \\
\text{awid} & \quad \text{‘go home’} & \quad \text{pag}\text{-}\text{awid}\text{-}e\text{n} & \quad \text{‘make someone go home’}
\end{align*}
\]

The circumfix \(\text{pag}\text{-}e\text{n}\) adds the meaning ‘make’ or ‘cause.’ It is called a causative.

For another example of circumfixes, see Textbox 4.1.

**Words may contain more than one prefix, more than one suffix, or combinations of both,** as in many of the Mohawk words seen earlier.

(10) Mohawk verb from (3)

\[
\text{niahhatakhenontie’}
\]

\[
\text{n-i-a-ha-ti-takhe-n-ontie’}
\]

\[
\text{there-thither-FACTUAL-they-all-run-to.there-along}
\]

\[
\text{‘they started running thither into it’}
\]

It is important to note that there are no infixes in this word. The morphemes \(n\) ‘there,’ \(i\) ‘thither,’ \(a\) FACTUAL, \(h\)a- 3 (THIRD PERSON), and \(ti\) PL (PLURAL) are all prefixes, because they all occur before the root. No morpheme appears inside of the root -\(\text{takhe}\) ‘run.’ Many languages, among them Turkish and the Eskimo-Aleut languages, contain only suffixes. Some others, among them Navajo and other Athabaskan languages, contain only prefixes.

**Words may also contain more than one root.** Such words are called compounds. English is particularly rich in compounds. English speakers frequently combine roots, even full words, to create new words, such as break-through, pig-pen, hot-dog, and many more. Frequent word-formation processes are said to be highly productive; this concept is further discussed in Textbox 4.2. Newer compounds are often still written as two words, even when they are pronounced and understood as a term for a single idea. The Mohawk example seen earlier in (2) is also a compound, built on the noun root -\(\text{nonhs}\) ‘house’ and the verb root -\(\text{ihsak}\) ‘seek’:

\[
\text{teni-nonhs-ihsak-ha ‘Let’s go house-hunting’}
\]
4.3 Finding morphemes: morphological analysis

How do we know what the morphemes are in a word? Speakers themselves are rarely conscious of morphological structure unless they have studied grammar. They obviously have some unconscious knowledge of morphology – we see this when they
create new words - but few could explain it. If we cannot ask speakers about morphology, how can we discover it in a language?

Morphological analysis is usually done by comparing sets of words, as we did for the Ilocano examples on the last several pages. If we find two similar words, one with a particular sequence of sounds and one without it, we can compare the meanings of the words and hypothesize that the additional sequence of sounds adds the additional meaning. Consider the following examples from English:

(11) English word pairs

happy cool
unhappy uncool

The meanings of the words with un- are the opposite of those without. We can hypothesize that English has a prefix un- meaning ‘not.’

Now examine the pairs of words below from Karo, a language of the Tupi-Guaraní family spoken in the Brazilian Amazon.

(12) Karo verbs (Gabas 1999)

kɔ ‘walk’
takɔ ‘make (someone) walk while walking alongside of him or her’
kɛt ‘sleep’
takɛt ‘put (someone) to sleep and sleep alongside him or her’
nɔğa ‘eat’
tanɔga ‘feed (someone) while eating with him or her’

Each of these verbs can occur with or without the element ta. The addition of ta to the verbs changes their meaning in a systematic way. It looks like we have a morpheme: a recurring sequence of sounds (ta) with a recurring meaning. As far as we can tell from the data in (12), this morpheme never appears on its own as a word: it is bound. Since it appears before roots (rather than after them or inside of them), it is a prefix. Semantically, it adds a sense of causation, causing someone to walk, causing someone to sleep, and causing someone to eat. It is thus a causative. But it adds a more specific meaning than the Ilocano causative we saw earlier. Here the causer is also participating in the caused activity: walking, sleeping, or eating. This Karo prefix ta- has been termed a comitative causative. These examples illustrate an important principle of morphological analysis. We cannot expect that every word in the free translation will correspond to a specific morpheme in the language under analysis. Skillful morphological analysis can require the ability to think beyond the translation to imagine what the actual meaning might be. (Linguists generally use the term gloss rather than “translation,” because they recognize that morphemes and words in one language...
do not always have perfect semantic equivalents in another. A gloss is simply the best approximation to the meaning. See Sidebar 4.7 for glossing conventions.

In conducting morphological analysis, we may not always find pairs of words like those in (11) and (12): one without an affix and one with it. Fortunately, there is another strategy for detecting morphemes. We may find sets of words that all contain a particular element, or a certain sequence of sounds. We can then compare the meanings of those words to see what they have in common. When a recurring form matches up with a recurring meaning, we can hypothesize that we have a morpheme with that form and that meaning.

Compare the Karo words in (13).

(13) Karo verbs (Gabas 1999)

\begin{verbatim}
oya'wan  'I left'
omåmnoy  'I saw myself'
owakán  'I am angry'
okət  'I walked'
\end{verbatim}

All of these words begin with o, though otherwise they are different. One involves leaving, one seeing, one being angry, and one walking. One is in the present, and the others are in the past. But all include the meaning ‘I.’ We can hypothesize that o- is a morpheme meaning ‘I.’

4.4 The meanings of morphemes

Do all languages express the same kinds of meanings in their morphemes? There certainly are similarities. Many languages have noun roots meaning ‘head,’ ‘fish,’ and ‘house,’ just like in English. Many have verb roots meaning ‘eat,’ ‘catch,’ and ‘kill.’

But languages also show differences in their inventories of roots. Often roots are simply not equivalent. Central Alaskan Yup’ik, an Eskimo-Aleut language spoken in southwestern Alaska, has roots for ‘head’ (nasqag) and ‘house’ (ca), like the languages in Table 4.2. But the Yup’ik root neqe: ‘fish’ also means ‘food’ and ‘eat.’ Furthermore, for many Yup’ik roots, there is simply no single equivalent English root. They can be translated only with a complex explanation. (The Yup’ik material in this chapter

SIDEBAR 4.7
GLOSSING CONVENTIONS
We write the gloss of a morpheme in small caps if it is a grammatical term like PLURAL (PL), NEGATIVE (NEG), CAUSATIVE (CAUS), or NOMINALIZER (NMLZ). If it is a translation, we write it between single quotation marks, like for the Mohawk suffix -ontie ‘along.’ Material in the target language, that is, the language under study, is usually given in italics, or, when handwritten, underlined.

For this chapter (and all subsequent chapters and language profiles) of this book, a full list of the glossing conventions used in each chapter can be found at the end of the chapter, just prior to the Exercises.

SIDEBAR 4.8
TRANSCRIPTION NOTE
Karo symbol IPA Phonetic description
’ (apostrophe) [ʔ] glottal stop

80 Marianne Mithun
Morphology: what's in a word?

### TABLE 4.2 Roots in several different languages

#### Some noun roots

<table>
<thead>
<tr>
<th>English</th>
<th>head</th>
<th>fish</th>
<th>House</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>Kopf</td>
<td>Fisch</td>
<td>Haus</td>
</tr>
<tr>
<td>French</td>
<td>tête</td>
<td>poisson</td>
<td>Maison</td>
</tr>
<tr>
<td>Engenni</td>
<td>ātōmū</td>
<td>ēsēnī</td>
<td>Ōmu</td>
</tr>
<tr>
<td>Mohawk</td>
<td>-nontsi</td>
<td>-itsu-</td>
<td>-nonhs-</td>
</tr>
<tr>
<td>Ilocano</td>
<td>ulo</td>
<td>ikān</td>
<td>Bal</td>
</tr>
<tr>
<td>Karo</td>
<td>nakd</td>
<td>ip</td>
<td>ka’a</td>
</tr>
</tbody>
</table>

#### Some verb roots

<table>
<thead>
<tr>
<th>English</th>
<th>eat</th>
<th>catch</th>
<th>Kill</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>ess-</td>
<td>fang-</td>
<td>töt-</td>
</tr>
<tr>
<td>French</td>
<td>mang-</td>
<td>attrapp-</td>
<td>tu-</td>
</tr>
<tr>
<td>Engenni</td>
<td>dhi</td>
<td>kunu</td>
<td>Čbēi</td>
</tr>
<tr>
<td>Mohawk</td>
<td>-k</td>
<td>-iena</td>
<td>-rio</td>
</tr>
<tr>
<td>Ilocano</td>
<td>koan</td>
<td>tiliw</td>
<td>Patay</td>
</tr>
<tr>
<td>Karo</td>
<td>to</td>
<td>‘iy</td>
<td>wi</td>
</tr>
</tbody>
</table>

comes from the speech of Elizabeth Charles, Elena Charles, and George Charles, and from Jacobson (1985).

(14) Some Yup'ik roots

keniq ‘front part of a parka cover, gathered up and used as a means of carrying things’

ella ‘world, outdoors, weather, universe, awareness, sense’

pay’u ‘to have one's legs so cramped by cold that one cannot move’

gapiar- ‘to skin a seal or other animal starting from the head and pulling the skin back over the body, rather than splitting the skin’

taarri- ‘to swat oneself or another in a steambath to tone muscles and stimulate sweating’

caqvir- ‘having shoes on the wrong feet’

narunte- ‘to act against accepted standards of behavior’
When we compare the meanings of affixes across languages, we find the same kinds of similarities and differences. Some affix meanings show up in language after language. For example, many languages have past-tense affixes, comparable to English -ed. Many have plural affixes, comparable to English -s. Many have nominalizers, like the Ilocano suffix -an we saw in giling-an 'grind-er' and the English suffix -er. Many have causatives like the Ilocano circumfix in pag-ragsak-en 'make someone happy.'

As seen in Textbox 4.3, affix meanings can vary across languages. Look at the meanings of the Yup’ik suffixes in (15).

**TEXTBOX 4.3 THE MEANINGS OF AFFIXES**

Although different languages can have similar affix categories, languages exhibit interesting differences. Some languages have no affixes that mark past tense, while others have elaborate sets of past-tense affixes, distinguishing degrees of remoteness, such as ‘immediate past,’ ‘recent past,’ ‘remote past,’ and ‘mythic past.’ Karo has a basic causative prefix ma-, as in ket ‘sleep,’ ma-ket ‘make someone sleep’; kot ‘walk,’ ma-kot ‘make someone walk’; capit ‘be fat,’ ma-capit ‘make someone fat.’ But as we saw, it also has a more specialized kind of causative prefix ta- meaning ‘cause and participate in the action.’ Yup’ik has an unusually rich array of causative suffixes, which differ in sometimes subtle ways. They add meanings such as ‘let,’ ‘allow,’ ‘permit,’ ‘cause,’ ‘compel,’ ‘force,’ ‘deliberately or intentionally cause,’ ‘try to cause,’ ‘tend to cause,’ and ‘wait for’ (as in ‘wait for something to boil’).

**(15) Some Yup’ik suffixes**

<table>
<thead>
<tr>
<th>Affix (yup’ik)</th>
<th>Meaning</th>
<th>Affix (yup’ik)</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>-kuaq</td>
<td>‘leftover’</td>
<td>arucetaar</td>
<td>‘dried fish skin’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>arucetaar-kuaq</td>
<td>‘uneaten, leftover dried fish skin’</td>
</tr>
<tr>
<td>-kuar</td>
<td>‘go by way of’</td>
<td>imarpig</td>
<td>‘sea’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>imarpig-kuar</td>
<td>‘sea’</td>
</tr>
<tr>
<td>-liqe</td>
<td>‘be afflicted in’</td>
<td>ilu-</td>
<td>‘inside, digestive track’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>liu-liqe</td>
<td>‘have a stomach-ache’</td>
</tr>
<tr>
<td>-illiqe</td>
<td>‘suffer from lack of’</td>
<td>mura-illiqe-</td>
<td>‘suffer from the lack of wood’</td>
</tr>
<tr>
<td>-ir</td>
<td>‘have cold’</td>
<td>it’ga-</td>
<td>‘foot’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>it’ga-ir-</td>
<td>‘have cold feet’</td>
</tr>
<tr>
<td>-taq</td>
<td>‘caught object’</td>
<td>pi</td>
<td>‘thing’</td>
</tr>
<tr>
<td>-tar</td>
<td>‘gather from nature’</td>
<td>mera-taq</td>
<td>‘caught animal or bird’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mer</td>
<td>‘water’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mer-tar-</td>
<td>‘fetch water’</td>
</tr>
</tbody>
</table>

The verb mer-tar- is used, for example, for getting water from a river. For getting water from a faucet or barrel, a different suffix, -ssaag, is used. **The affixes that develop in languages are no accident. They grow out of distinctions**
that speakers have chosen to express most often in daily speech, over generations, centuries, even millennia. Some affixes reflect concepts that are important to human beings all over the world, like causation. Others reflect the environmental and cultural concerns of individual societies. Textbox 4.5 discusses the meanings expressed by roots and affixes.
4.5 The shapes of morphemes

Once we identify a morpheme, can we be confident that it will always show up with the same form? Let’s consider English plurals. To form a plural, it looks like we just add -s.

(18) English plurals

newt newt-s
skink skink-s
toad toad-s
frog frog-s
tadpole tadpole-s
chameleon chameleon-s
salamander salamander-s
iguana iguana-s
gecko gecko-s

But are all of these suffixes really the same? If you pronounce the plural words carefully, you’ll notice that they do not all end in the same sound. Some end in voiceless [s], and some end in voiced [z]. Pronounce the plural nouns in (18), checking whether you can hear the difference. Are these just allophones in English, or are they separate phonemes? To find out, we look for minimal pairs. There are plenty of them: sip and zip are different words, for example, with different meanings. The sounds /s/ and /z/ are thus separate phonemes in English.

If English speakers hear a new noun, they somehow know which form to use for the plural. What word would you use to talk about more than one swip? How about more than one grib? Speakers automatically use the voiceless /s/ with swips, and the voiced /z/ with gribs, though they have never heard these words before.

As we saw in Chapter 3, if we look carefully at where each form occurs, we see a pattern. The voiceless plural /s/ comes after nouns ending in p, t, and k. The voiced plural /z/ comes after nouns ending in d, g, l, n, r, w, y, and vowels. We can make a generalization about these contexts: the form /s/ occurs only after voiceless sounds, and the form /z/ occurs only after voiced sounds. The alternate forms of a morpheme, like the /s/ and /z/ plural here, are called allomorphs. (Note that the plural of allomorph uses the voiceless /s/, because this word ends in the voiceless sound /f/.) If we find that a morpheme has allomorphs, we need to mention this in our description of the language, specifying what the various forms are and where they occur. For our plural allomorphs, we have two forms so far.

(19) -s PLURAL
     s used after voiceless sounds
     z used after voiced sounds
A description like that in (19) is termed **item and arrangement**: we specify how the items are arranged, that is, where each allomorph occurs.

Allomorphy can also be described in another style, termed **item and process**, using symbols we saw in the chapter on phonology (see Sidebar 4.10). We want to say that the plural suffix becomes voiced after a voiced sound. We begin with the basic sound, here s. We then add an arrow → for ‘becomes’ or ‘is pronounced.’ To the right of the arrow, we specify how it changed (in this case by becoming voiced). Next is a slash /, which announces that the context is coming up. The blank ___ shows us where the sound in question fits into the context.

(20) -s PLURAL
-s → [+voiced] / [+voiced] ___
-s becomes voiced in the context after any voiced sound

Are -s and -z the only allomorphs of the English plural? What about the plural of grouse? How about thrush and finch? How about partridge? Even if you can’t recognize a thrush, you probably know that its plural would be pronounced with [az]. We find the same plural with **grouse, finch, and partridge**. The voicing rule we worked out in (20) doesn’t seem to be working here. Are these erratic exceptions? No, we just need an additional rule. The nouns **grouse, thrush, finch, and partridge** all end in hissing fricatives or affricates called **stridents** or **sibilants**. This, too, is systematic. Our new analysis might look something like this.

(21) -s PLURAL
-s → az / [+strident] ___
-s → [+voiced] / [+voiced] ___

Anyone who has learned another language after childhood knows that allomorphs can make the task harder. Why would a language have them? Sometimes the answer is hidden in the history of the language, but this time, we can see some reasons. When speakers already have their vocal folds vibrating at the end of a noun (a noun ending in a voiced sound like b, d, g, w, l, n, r, a, e, o, etc.), it would take extra work to interrupt the vibration and cut off the voicing in order to produce a voiceless /s/. It is easier to just let the vibration continue to the end of the word. This allomorphy allows speakers to expend less effort. But what about the other allomorph, the /az/ we find after stridents? If we pronounced the plural of a noun like **grouse** with just /s/, it would be nearly impossible to hear the suffix: [graʊs]. The added vowel helps listeners hear the plural. These are two of the many forces that are always at work shaping the development of languages: ease for speakers and clarity for hearers.

Now what about the plurals **cactus/cacti**, **phenomenon/phenomena**, and **cherub/cherubim**? Are these just random irregularities? Not really. They are indeed exceptions
to our rules, but there are reasons they exist. These nouns were brought into English from other languages, specifically Latin, Greek, and Hebrew. The English speakers who originally adopted them knew enough about the donor languages to bring their plural forms along with them. The suffix -i in *cact-i* is a Latin plural, the suffix -a is a Greek plural, and the suffix -im is a Hebrew plural. When enough nouns have been brought into one language from another, and their plural forms have been brought in along with them, speakers may begin to discern this pattern and apply it to new forms they encounter with the same endings. If English speakers want to pluralize a word ending in -*us* (which happens to be a neuter singular ending in Latin), they may decide that it should follow the same pattern as *cactus* and pluralize it with -i, even if they have not heard its plural form and don’t know whether it came from Latin. Sidebar 4.11 further explains how such patterns may come to be regularized.

We have seen two kinds of allomorphy. The plural allomorphs -s, -z, and -ǝz are said to be **phonologically conditioned.** The choice of which allomorph to use depends simply on the sound before it: -ǝz is used after stridents, -s is used after other voiceless sounds, and -z is used after other voiced sounds. The plural allomorphs -i, -a, and -im are said to be **lexically conditioned.** One simply has to know the word (lexeme) it is attached to in order to know the plural.

Now let’s look at allomorphy in another language. Ilocano has a prefix that can form verbs meaning ‘move an object to the location specified by the root.’

(22) **Ilocano verbalizer**

<table>
<thead>
<tr>
<th>English</th>
<th>Ilocano verbalizer</th>
<th>Ilocano meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>sakmol</em></td>
<td>i-sakmol</td>
<td>‘put something into the mouth’</td>
</tr>
<tr>
<td><em>ruar</em></td>
<td>i-ruar</td>
<td>‘take something outside’</td>
</tr>
<tr>
<td><em>ditoy</em></td>
<td>i-ditoy</td>
<td>‘put something here’</td>
</tr>
<tr>
<td><em>ngato</em></td>
<td>i-ngato</td>
<td>‘put something up’</td>
</tr>
<tr>
<td><em>baba</em></td>
<td>i-baba</td>
<td>‘put something down’</td>
</tr>
<tr>
<td><em>abut</em></td>
<td>y-abut</td>
<td>‘put something in a hole’</td>
</tr>
<tr>
<td><em>uneg</em></td>
<td>y-uneg</td>
<td>‘put something inside’</td>
</tr>
<tr>
<td><em>abay</em></td>
<td>y-abay</td>
<td>‘place beside, compare’</td>
</tr>
</tbody>
</table>

Ilocano seems to have a **verbalizer** prefix *i:-* it turns words into verbs. But when we reach the pair *abut* ‘hole,’ *yabut* ‘put something in a hole,’ we see a change in the shape of the prefix. Here the prefix is *y.* (The letter *y* in the Ilocano spelling system represents a palatal glide, IPA [j]). The forms *i-* and *y-* appear in the same
position in the word, immediately before the root, and they contribute the same meaning, ‘put something somewhere.’ These two facts together suggest that we have one morpheme with two allomorphs, that is, two forms of the same morpheme. Sidebar 4.12 illustrates one of the ways in which allomorphy can first arise.

When it appears that we have two (or more) allomorphs, we look to see where each one occurs. When we have phonologically conditioned allomorphs, the distribution of allomorphs is usually determined by the sounds directly preceding or following them. Since there is nothing before our prefix here, the most promising place to look is after it. We find:

_i- before s, r, d, ng, b
_y- before a, u

(Recall that ng is a digraph representing the single consonant [ŋ].) Can we make any generalizations about the context, as we did for the English plural? The _i- occurs only before consonants, and the _y- only before vowels. This situation is called complementary distribution. In Chapter 3, we saw allophones of a single phoneme in complementary distribution; we now see that allomorphs of a single morpheme can also be in this relationship. The two forms of our prefix never occur in the same context, a confirmation that we do have allomorphs. We can now describe our prefix with its allomorphs.

(23) _i- VERBALIZER
   i→ y / __ V

This rule says that the sound _i is pronounced _y before a vowel.

Note that roots can have allomorphs as well. Compare _leaf_ and _leaves_: /lif/ ~ /liv/.

### 4.6 Are affixes always pieces of words?

Consider the English nouns _mouse/mice_ and _man/men_. These plurals seem irregular: they don’t match our rules. But they are not random. They are the result of a sequence of developments in the history of the English language, each of which made sense at the time. But how do we describe these plurals today? Do these forms have a plural affix? The answer is yes, but this affix has a different character from the prefixes, suffixes, infixes, and circumfixes we’ve seen. Here the plural affix could not be described as a piece of a word that is added, but as a change internal to the root. **Sometimes affixes are simply changes in a form, not necessarily pieces added.** This is true of the plurals of _mouse_ and _man_ in both, plurality is signaled by a change in the vowel. It is also true of the past tense on verbs like _see/saw_ and _run/ran_.

---

**SIDEBAR 4.12 ORIGINS OF ALLOMORPHY**

Try pronouncing the sequence _i-abay_ several times quickly. You will note that with speed, the vowel naturally turns into a glide. This is one of the ways that allomorphs can develop in languages over time.
This allomorphy is lexically conditioned, because these changes could not be predicted on the basis of the sounds of the roots alone.

Now compare the Engenni verbs in (24). The marks over the vowels indicate tone. The absence of a mark indicates high tone, and an acute accent (è) indicates low tone.

(24) Engenni (Thomas 1978)

dìre ‘cook’
dìre ‘cooks’
dìre ‘will cook’

The verb dìre ‘cook’ alone has high tone on both syllables (HH). The word dìre ‘cooks’ has low tone on the first syllable and high tone on the second (LH). The future form dìre ‘will cook’ has high tone on the first syllable and low on the second (HL). Here, too, it is easier to describe the tenses as changes made to the root than as pieces of words. The present-tense affix is a LH tone structure, while the future-tense affix is a HL structure.

4.7 Words, lexemes, and lexicalization

At the beginning of this chapter, we considered how to identify words in an unfamiliar language. We can now take this question a bit further. Consider the English words lizard and lizards. We know that they are different words. But we would not expect to find both in a dictionary; we will find only lizard. The same is true of the words talk, talked, and talking. They are all different words, but we will find only talk in the dictionary. Words like lizard and lizards are said to belong to the same lexeme. Similarly, talk, talked, and talking belong to the same lexeme. The plural suffix -s on lizard does not form a new vocabulary item; it simply gives us another form of the same one, used for referring to more than one of the same thing. A lexeme is a basic vocabulary item, something we might find as an entry in a dictionary (or in our mental lexicon).

Morphology that does not form a new lexeme, like the plural -s on nouns and -ed and -ing on verbs, is called inflection. Inflectional affixes never change the word class of the word that they attach to. Gavotte is a noun, and the plural gavotte-s is still a noun. Inflectional affixes tend to be highly productive. If we learn a new English count noun like gavotte, we expect that it will have a plural, though we may not always be able to predict the form of the plural. Inflectional affixes tend to contribute predictable meanings. If we know that the noun gavotte means ‘a medium-paced dance popular in the eighteenth century,’ we may never have heard it in the plural, but we can predict that gavotte-s will mean more than one gavotte. For examples of some inflectional affixes in English, see Textbox 4.6.

Languages differ in the meanings expressed in their inflectional morphology, though certain inflectional categories appear in language after language. Plural and
tense inflection are both common cross-linguistically. Another common kind of inflectional morphology is **case**. To see what case morphology can do, look at the Latvian sentences given in (25).

(25) Latvian (Uldis Balodis, p.c.)

a. Laukā dzīvo tārps.
   field live.3 worm
   ‘A worm lives in the field.’

b. Lauk ir liels.
   field COP.3 big
   ‘The field is big.’

c. Putns ēd tārpu.
   bird eat.3 worm
   ‘A bird is eating the worm.’

d. Putnā tagad dzīvo tārps.
   bird now live.3 worm
   ‘The worm now lives in the bird.’

e. Vīrs ēd putnu.
   man eat.3 bird
   ‘A man is eating the bird.’

f. Mežā dzīvo vilks.
   forest live.3 wolf
   ‘A wolf lives in the forest.’

g. Mežs ir baismīgs.
   forest COP.3 terrifying
   ‘The forest is terrifying.’

h. Vīlks ēd vīru.
   wolf eat.3 man
   ‘The wolf is eating the man.’

i. Vīlks tagad dzīvo tārps.
   wolf now live.3 worm
   ‘The worm now lives in the wolf.’

Each of the nouns occurs in several sentences. Now look more closely at the forms of these nouns. Start with the word for ‘field.’ In (25a) it is *laukā*, but in (25b) it is *lauks*. Look at the words for ‘bird.’ In (25c) it is *putns*, in (25d) it is *putnā*, and in (25e) it is *putnu*. Take a moment to see whether you can tell why the forms keep changing. The other nouns show similar changes.

Each Latvian noun here contains a suffix that identifies its role in the sentence. These suffixes are what are termed case-markers. The ending -s indicates that the noun is the subject.
The word putn-s 'bird' is the subject of (25c) ‘The bird is eating the worm.’ Morphemes that identify subjects, like this -s, are called nominative case-markers. The word putn-u 'bird' is the direct object of (25e): ‘The man is eating the bird.’ Morphemes that identify objects, like the -u, are called accusative case-markers. The word putn-ā specifies a location in (25d): ‘The worm lives in the bird.’ The ā is a locative case suffix. Case morphology occurs in languages all over the world, though certainly not in every language. Languages that do have case morphology show fascinating variation in the number and functions of their case categories. Latvian itself has other case categories in addition to those shown here. The origins of several of the Latvian words given here are discussed in Textbox 4.7.

**TEXTBOX 4.7 LATVIAN**

Look again at the Latvian words in (24). Do any look like words with similar meanings in other languages that you know? Latvian is spoken at the eastern end of the Baltic Sea, across from Sweden. It is remotely related to English: both are descended from the same ancestral language, Indo-European. Several of the Latvian words in (24) have English cognates that are words descended from the same Indo-European word. The Proto-Indo-European root for ‘wolf,’ for example, is reconstructed as *wilk*-,

nominative form *wlk*-os. The Latvian root is vilk-,

nominative vilk-s. The Proto-Indo-European verb root

*hydration- is the ancestor of Latin *ed-, Latvian *ed, and English eat. The Proto-Indo-European noun *ultrō- ‘free man’ developed into Latin vir and Latvian *vir- ‘man.’ Proto-Indo-European had case suffixes, which continued into Latin, Latvian, and other languages in the family. English, like its relatives, also once had case suffixes, but they have been lost.

Not all morphology is inflectional. Morphological processes that create new lexemes are derivational. From the root talk, a new lexeme was created in the twentieth century: *talk-ie* ‘a film with a soundtrack, as distinct from a silent film.’ This suffix -ie is not fully productive: there are many nouns that speakers would not add it to. Can you imagine a word window-ie? **Derivational suffixes do not necessarily add predictable meaning**. If you only knew the root talk, would you be able to predict the meaning of talkie? **Derivational morphology can change the word class of the lexeme it is applied to**, though this is not always the case. (See Textbox 4.8 for some examples of derivational affixes in English.) The nominalizers and verbalizers we saw earlier are derivational affixes. Added to the English verb govern, the derivational suffix -ment creates a new noun: government. Added to the Yup’ik noun imarpig ‘sea,’ the derivational suffix -kuar ‘go by way of’ creates a new verb: imarpig-kuar ‘go by sea.’

Speakers recognize the difference between those words that already exist in their language and those that could but have not yet been created. This knowledge is sometimes referred to as the mental lexicon. English speakers know that the words joyful and peaceful are part of English, and that giggleful and warful are not, or at least not yet. Words that do exist are called actual words. Words that do not yet exist but could, according to patterns already present in the language, are called potential words. When a speaker first coins a new word through derivation or compounding,
listeners will often recognize it as an innovation. At this point it is called a **nonce formation**. If other speakers pick it up and use it, it can become an accepted part of the vocabulary or lexicon of the language. This process, by which a morphological formation comes to be recognized as an established word in the language, is called **lexicalization**. A lexicalized word will be understood and learned as a unit, rather than processed as a sequence of individual morphemes. It will have a meaning of its own, which may or may not correspond to the sum of its parts. The word *joy-ful* means ‘full of joy’ (as we would expect), but *aw-ful* no longer means ‘full of awe’ (though it once did), and *cup-ful* does not mean ‘full of cups.’

### 4.8 Typology

Languages are often classified according to their structural characteristics. The classification of languages according to structural traits is a goal of **linguistic typology**, the study of language types. Perhaps the oldest and still most commonly cited typologies are based on morphological structure. Structures like that seen in the Engenni example at the beginning of this chapter, with words that generally consist of just one morpheme, are called **isolating** or **analytic**. Structures like those in Mohawk are called **polysynthetic**, because words tend to be made up of many (poly-) parts put together (synthetic). This typological dimension, roughly the average number of morphemes per word, is called the degree of synthesis. Languages like English and Japanese are often described as “mildly synthetic”, that is, their words may consist of more than one morpheme, but they do not generally show the elaborate morphological complexity of languages like Mohawk.

A second frequently cited typological feature is the degree of fusion. This characteristic pertains to the clarity of divisions between morphemes in a word. Morphological structures with sharp boundaries between morphemes are described as **agglutinat-ing**. The English word *understandable* shows agglutinating structure: *under-stand-able.*

---

**Textbox 4.8 SOME DERIVATIONAL MORPHEMES IN ENGLISH**

<table>
<thead>
<tr>
<th>Change the part of speech</th>
<th>Do not change the part of speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ing</td>
<td>verb to noun</td>
</tr>
<tr>
<td>-ly</td>
<td>adjective to adverb</td>
</tr>
<tr>
<td>-er, -or</td>
<td>noun to verb</td>
</tr>
<tr>
<td>-en</td>
<td>adjective to verb</td>
</tr>
<tr>
<td>-able</td>
<td>verb to adjective</td>
</tr>
<tr>
<td>-ment</td>
<td>verb to noun</td>
</tr>
<tr>
<td>-less</td>
<td>without; noun to adjective</td>
</tr>
<tr>
<td>en-</td>
<td>noun or adjective to verb</td>
</tr>
</tbody>
</table>

**SOME DERIVATIONAL MORPHEMES IN ENGLISH**

<table>
<thead>
<tr>
<th>Change the part of speech</th>
<th>Do not change the part of speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ing</td>
<td>verb to noun</td>
</tr>
<tr>
<td>-ly</td>
<td>adjective to adverb</td>
</tr>
<tr>
<td>-er, -or</td>
<td>noun to verb</td>
</tr>
<tr>
<td>-en</td>
<td>adjective to verb</td>
</tr>
<tr>
<td>-able</td>
<td>verb to adjective</td>
</tr>
<tr>
<td>-ment</td>
<td>verb to noun</td>
</tr>
<tr>
<td>-less</td>
<td>without; noun to adjective</td>
</tr>
<tr>
<td>en-</td>
<td>noun or adjective to verb</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change the part of speech</th>
<th>Do not change the part of speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ing</td>
<td>verb to noun</td>
</tr>
<tr>
<td>-ly</td>
<td>adjective to adverb</td>
</tr>
<tr>
<td>-er, -or</td>
<td>noun to verb</td>
</tr>
<tr>
<td>-en</td>
<td>adjective to verb</td>
</tr>
<tr>
<td>-able</td>
<td>verb to adjective</td>
</tr>
<tr>
<td>-ment</td>
<td>verb to noun</td>
</tr>
<tr>
<td>-less</td>
<td>without; noun to adjective</td>
</tr>
<tr>
<td>en-</td>
<td>noun or adjective to verb</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change the part of speech</th>
<th>Do not change the part of speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ing</td>
<td>verb to noun</td>
</tr>
<tr>
<td>-ly</td>
<td>adjective to adverb</td>
</tr>
<tr>
<td>-er, -or</td>
<td>noun to verb</td>
</tr>
<tr>
<td>-en</td>
<td>adjective to verb</td>
</tr>
<tr>
<td>-able</td>
<td>verb to adjective</td>
</tr>
<tr>
<td>-ment</td>
<td>verb to noun</td>
</tr>
<tr>
<td>-less</td>
<td>without; noun to adjective</td>
</tr>
<tr>
<td>en-</td>
<td>noun or adjective to verb</td>
</tr>
</tbody>
</table>

These English derivational affixes can be contrasted with the English inflectional affixes in Textbox 4.6.
Morphological structures in which boundaries are not clear are described as fusional. The English word *men* consists of the root *man* plus plural marking, but it is difficult to pull the two apart. A language may have both agglutinating and fusional structures, like English, but languages often show general tendencies one way or the other. Quechua, a language of Peru, Bolivia, and Chile (one variety is the subject of a Language Profile) is often cited as an agglutinating language. The second line of the South Conchucos Quechua example below shows the individual morphemes in the word. The top line shows how the word is pronounced. The individual shapes of each morpheme remain unchanged.

(26) Quechua (Hintz 2008)
Tsypitakuti-tsiya-muru-usi
tsay-pita-kuti-tya-mu-ru::si
return-CAUS-PL-far-PST-I-even
‘I made them return.’
Now compare the Quechua example to the word below in Seneca, an Iroquoian language of present western New York State. (Seneca is also the subject of a Language Profile.) Seneca is more fusional. Again look first at the second line, which shows the shape of each morpheme. Compare this to the top line, which shows what happens when they are combined in a word.

(27) Seneca (Wallace Chafe, p.c.)
Ögí:wanâ:go:
wa’-wak-rihw-a-nehrako-
FACTUAL-me-matter-LINKER-surprise-PFV
‘It surprised me.’

All of the morphemes listed in the second line are present, but a series of historical sound changes has obscured their boundaries that only those steeped in the history of Iroquoian languages can now decipher them.

4.9 Morphology, language, and us

Morphology is connected to both phonology and syntax, but is distinct from each. Some of the patterns we see in phonology have echoes in morphology: just as phonemes can take different forms (allophones) in different contexts, so, too, can morphemes can take different forms (allomorphs) in different contexts (for a discussion of how children learn these forms and contexts, see Textbox 4.10).

Few parents explain affixes and allomorphy to their children. You have probably never heard a parent say: “To make a plural, dear, you must add the suffix -es after sibilants, -s after other voiceless sounds, and -z after other voiced sounds, unless the word requires a special, lexically-conditioned allomorph.” Language acquisition occurs naturally, without explicit instruction. In acquiring language, children often extend regular patterns beyond where they are found in adult speech. For example, children often say mouses and mans as the plurals of mouse and man. It is clear that the child has learned the regular rule of plural formation and is applying it to all nouns. In time, the child will learn the lexically conditioned forms of the plural and will automatically substitute them for the forms created with the regular plural affix for certain words.

Both morphology and syntax consist of patterns of combining smaller pieces of language to form larger ones: morphemes are combined into words, and words are combined into sentences. But these two levels of structure are not the same. Speakers do not usually exploit their knowledge of morphological patterns, particularly derivational ones, every time they speak. They typically use this knowledge only on certain occasions to invent vocabulary, when they feel the need for a new word.
Listeners often notice when someone has created a new word, even if the parts are familiar. If I told you my skirt was *unslitable*, you might or might not understand that I cannot make a slit in it, but you would probably realize that you had never heard the term. By contrast, speakers often form new sentences as they speak, on the basis of their knowledge of syntactic patterns. If I told you, *I really want a persimmon-striped umbrella*, your first reaction would probably not be *Gee, I’ve never heard that sentence before*.

Comparing morphological patterns in different languages reveals certain features characteristic of all human language. Languages are shaped by human cognitive abilities, such as pattern recognition and extension, memory, and the ability to routinize recurring tasks. It is in fact this deeper cognitive similarity that can lead to some of the most interesting differences we see across languages, as speakers automate frequent patterns of expression. Modern morphological distinctions are largely the result of what speakers have chosen to express the most frequently over the course of development of their languages.

**CHAPTER SUMMARY**

This chapter has provided a glimpse into the kinds of structures inside words in different languages. We began with the question of how to identify words in an unfamiliar language. The elements that make up words are called morphemes. Morphemes can be classified as bound or free, roots or affixes, prefixes, suffixes, infixes, or circumfixes, and inflectional or derivational. Morphological structure can be discovered in an unfamiliar language, and we can identify and describe alternating forms of morphemes, or allomorphs. Languages can vary widely in their morphological structure: in how much information is typically carried within a word (degree of synthesis) and how clear the boundaries between elements are (degree of fusion). They can also differ in fascinating ways in terms of the kinds of meanings expressed by morphemes. Such differences are not accidental; they grow out of speakers’ use of language for everyday communication, and we can see this reflected in how certain words have developed over time. This chapter has begun to reveal the remarkable diversity of linguistic structures; information is packaged and presented in very different ways across languages.

Morphemes are the building blocks of words, and words, in turn, are the building blocks of sentences. How words combine into sentences and the grammatical relationships between words are the central concerns of the field of syntax. But first we’ll need to explore in more depth the different kinds of words that one finds in language, and the behaviors and characteristics that allow us to identify word classes.
SUGGESTIONS FOR FURTHER READING


This is a book-length, reader-friendly introduction to morphology, aimed at an audience without prior experience of the field. It covers all of the major topics: wordhood, word classes, inflection and derivation, relations between morphology and phonology, morphology and syntax, morphology and semantics, and productivity.


Chapter 2 (on morphological typology) and Chapter 5 (on noun and noun-phrase operations) are especially relevant.


This little book is a classic, presenting ideas that have captivated linguists for generations. It is well worth reading the whole book through. Chapter 4 discusses morphological structure: compounding, prefixing, suffixing, infixing, internal vocalic change, consonantal change, reduplication, and pitch.

EXERCISES

1. Ainu is a language isolate (a language with no known relatives) indigenous to Northern Japan and adjacent Russia. (Data here are drawn from Refsing 1986: 134.)

   a. wen ‘be bad’ wenno ‘badly’
   b. pirka ‘be fine, be good’ pirkano ‘well’
   c. esikin ‘be kind, friendly’ esikino ‘kindly’
   d. araskay ‘be very able’ araskayno ‘successfully’
   e. asin ‘be new’ asinno ‘newly’
Like many other languages, Ainu does not have a separate adjective category. Qualities are expressed by verbs, like those in the left column above. Do a complete morphological analysis of the data here.

i. For each lexical category (verb, noun, etc.), provide a list of roots in *italics* with their glosses in ‘*single quotes.*’

ii. For each category, provide a list of affixes (again in *italics*, with appropriate hyphens) with glosses in ‘*single quotes*’ for translations, SMALL CAPS for grammatical terms. If you can think of a single-word gloss for your affix(es), use that. If not, provide an explanation of the meaning.

SIDEBAR 4.16
TRANSCRIPTION NOTE

<table>
<thead>
<tr>
<th>Ainu symbol</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>postalveolar lax plosive, sometimes approaching [d] but [ɾ] between vowels</td>
</tr>
</tbody>
</table>

2. The Carib language, a member of the Cariban family, is spoken over the entire coastal area of Guiana. Other members of this language family are found all over northern South America, from Columbia to the mouth of the Amazon and from the coast of Guiana to far into the south and southwest of Brazil. (Data here are drawn from Hoff 1968.)

SIDEBAR 4.17
TRANSCRIPTION NOTE

<table>
<thead>
<tr>
<th>Carib symbol</th>
<th>IPA</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>:</td>
<td>[:]</td>
<td>high back unrounded vowel</td>
</tr>
<tr>
<td>y</td>
<td>[j]</td>
<td>high back unrounded vowel</td>
</tr>
<tr>
<td>꜉</td>
<td>[u]</td>
<td>high back unrounded vowel</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Carib</th>
<th>IPA</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. enë</td>
<td>‘see’</td>
<td></td>
</tr>
<tr>
<td>b. wose:ne</td>
<td>‘look at each other’</td>
<td></td>
</tr>
<tr>
<td>c. enepotï</td>
<td>‘see repeatedy’</td>
<td></td>
</tr>
<tr>
<td>d. we:i</td>
<td>‘become’</td>
<td></td>
</tr>
<tr>
<td>e. weipotï</td>
<td>‘become repeatedly’</td>
<td></td>
</tr>
<tr>
<td>f. uxku</td>
<td>‘try’</td>
<td></td>
</tr>
<tr>
<td>g. u:kapotï</td>
<td>‘try again and again’</td>
<td></td>
</tr>
<tr>
<td>h. w:se:ne:po:tï</td>
<td>‘see each other all the time’</td>
<td></td>
</tr>
<tr>
<td>i. en:kepi</td>
<td>‘see no longer’</td>
<td></td>
</tr>
<tr>
<td>j. se:ka:po:tï</td>
<td>‘tear completely to shreds’</td>
<td></td>
</tr>
<tr>
<td>k. wo:mï</td>
<td>‘go in all together’</td>
<td></td>
</tr>
<tr>
<td>l. ema:mï</td>
<td>‘see day breaking’</td>
<td></td>
</tr>
<tr>
<td>m. enaimi:ma</td>
<td>‘run’</td>
<td></td>
</tr>
<tr>
<td>n. enai:ma</td>
<td>‘run out of’</td>
<td></td>
</tr>
<tr>
<td>o. sek:apotï</td>
<td>‘tear to shreds’</td>
<td></td>
</tr>
<tr>
<td>p. enë:ri</td>
<td>‘drink’</td>
<td></td>
</tr>
<tr>
<td>q. akisma</td>
<td>‘tease’</td>
<td></td>
</tr>
<tr>
<td>r. ena:mì:ma</td>
<td>‘see dawn turning into daylight’</td>
<td></td>
</tr>
<tr>
<td>s. wi:to</td>
<td>‘go’</td>
<td></td>
</tr>
<tr>
<td>t. wo:mì</td>
<td>‘go in’</td>
<td></td>
</tr>
<tr>
<td>u. akis:kepi</td>
<td>‘stop teasing’</td>
<td></td>
</tr>
<tr>
<td>v. i:ma:tu:ma</td>
<td>‘run out of completely’</td>
<td></td>
</tr>
<tr>
<td>w. enë:ri:ma</td>
<td>‘drink up’</td>
<td></td>
</tr>
<tr>
<td>x. se:kà</td>
<td>‘tear’</td>
<td></td>
</tr>
<tr>
<td>y. wi:to:po:tï</td>
<td>‘go repeatedly’</td>
<td></td>
</tr>
</tbody>
</table>
Morphology: what’s in a word?

Do a complete morphological analysis of the data here.

i. For each lexical category (verb, noun, etc.), provide a list of roots in *italics* with their glosses in ‘single quotes.’

ii. For each category, provide a list of affixes (again in *italics*, with appropriate hyphens) with glosses in ‘single quotes’ for translations, SMALL CAPS for grammatical terms. Different types of affixes (prefixes, suffixes, etc.) should be given in separate lists. If you can think of a single-word gloss for your affix(es), use that. If not, provide an explanation of the meaning.

For the purposes of this problem, ignore vowel length.

3. Daga is spoken in the Owen Stanley Mountains of the Central District of Papua New Guinea. (Data for this problem were drawn from: Murane 1974.)

   a. mamana ‘my father’
   b. inaga ‘your mother’
   c. yame ‘his eye(s)’
   d. nanimu ‘their hands’
   e. goanaya ‘your livers’
   f. pusinu ‘our feet’
   g. noga ‘your mouth’
   h. inana ‘my mother’
   i. done ‘its horn’
   j. pusina ‘my foot, my feet’
   k. yamga ‘your eye(s)’
   l. evene ‘his friend’
   m. inase ‘its tracks’
   n. goanana ‘my liver’
   o. nonu ‘our mouths’
   p. yamu ‘their eyes’
   q. noya ‘your mouths’
   r. naniga ‘your hand(s)’
   s. yame ‘her eye(s)’
   t. inasu ‘our mother’
   u. mamamu ‘their father’
   v. inaya ‘your mother’
   w. goaninu ‘our younger sibling’
   x. tase ‘her older sibling’
   y. mamanu ‘our father’
   z. yamya ‘your eyes’

SIDEBAR 4.18
TRANSCRIPTION NOTE
Daga symbol | IPA
---|---
y | [j]

Do a complete morphological analysis of the data here.

i. For each lexical category (verb, noun, etc.), provide a list of roots in *italics* with their glosses in ‘single quotes.’

ii. For each category, provide a list of affixes (again in *italics*, with appropriate hyphens) with glosses in ‘single quotes’ for translations, SMALL CAPS for grammatical terms. Different types of affixes (prefixes, suffixes, etc.) should be given in separate lists. If you can think of a single-word gloss for your affix(es), use that. If not, provide an explanation of the meaning.

iii. Do you have any ideas about whether the roots here are bound or free?
4. Lezgien is a language of the Nakh-Daghestanian family spoken by about 400,000 people in southern Daghestan and northern Azerbaijan in the eastern Caucasus. (Data here are drawn from Haspelmath 1993: 107.)

Do a complete morphological analysis of the data here.

i. For each lexical category (verb, noun, etc.), provide a list of roots in italics with their glosses in single quotes.

ii. For each category, provide a list of affixes (in italics, with appropriate hyphens) with glosses in single quotes for translations, SMALL CAPS for grammatical terms.

5. Samala, also known as Ineseño Chumash, is a Native California language, indigenous to the Santa Inez Valley near Santa Barbara, on the South Coast. It is a member of the Chumashan language family. (Data here are drawn from Applegate 1998.)

SIDEBAR 4.19
TRANSCRIPTION NOTE

<table>
<thead>
<tr>
<th>Lezgian symbol</th>
<th>IPA</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>č</td>
<td>[tʃ]</td>
<td>voiceless uvular fricative</td>
</tr>
<tr>
<td>š</td>
<td>[ʃ]</td>
<td>voiceless uvular fricative</td>
</tr>
<tr>
<td>x</td>
<td>[χ]</td>
<td>voiceless uvular fricative</td>
</tr>
<tr>
<td>ü</td>
<td>[y]</td>
<td>velar ejective</td>
</tr>
<tr>
<td>k'</td>
<td>[k']</td>
<td>velar ejective</td>
</tr>
</tbody>
</table>

Do a complete morphological analysis of the data here.

i. For each lexical category (verb, noun, etc.), provide a list of roots in italics with their glosses in ‘single quotes.’

ii. For each category, provide a list of affixes (in italics, with appropriate hyphens) with glosses in ‘single quotes’ for translations, SMALL CAPS for grammatical terms.

SIDEBAR 4.20
TRANSCRIPTION NOTE

<table>
<thead>
<tr>
<th>Chumash symbol</th>
<th>IPA</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>[ts]</td>
<td>voiceless alveolar affricate</td>
</tr>
<tr>
<td>p'</td>
<td>[p']</td>
<td>ejective consonant</td>
</tr>
<tr>
<td>[ʔ]</td>
<td>[ʔ]</td>
<td>glottal stop</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chumash symbol</th>
<th>IPA</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>[ts]</td>
<td>voiceless alveolar affricate</td>
</tr>
<tr>
<td>p'</td>
<td>[p']</td>
<td>ejective consonant</td>
</tr>
<tr>
<td>[ʔ]</td>
<td>[ʔ]</td>
<td>glottal stop</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chumash symbol</th>
<th>IPA</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>[ts]</td>
<td>voiceless alveolar affricate</td>
</tr>
<tr>
<td>p'</td>
<td>[p']</td>
<td>ejective consonant</td>
</tr>
<tr>
<td>[ʔ]</td>
<td>[ʔ]</td>
<td>glottal stop</td>
</tr>
</tbody>
</table>

a. cajxana ‘teahouse’
b. ktabxana ‘library’
c. küxban ‘idler’
d. qul ‘kitchen gardens’
e. qul ‘bird’
f. sal ‘kitchen garden’
g. wak ‘pigs’
h. caj ‘tea’
i. qul ‘birds’

j. qul ‘bird’
k. cajxana ‘printing plant’
l. ktab ‘book’
m. qul ‘brothel’
n. qul ‘herd of cattle’
o. caj ‘printing’
p. qul ‘prostitute’
q. qul ‘cattle herder’
r. wak ‘swineherd’
s. qul ‘kitchen gardener’
t. küx ‘street’
u. caj ‘tea’
w. qul ‘poultry farmer’
x. cajxana ‘printing plant’
y. caj ‘printing’
z. qul ‘prostitute’
a. wak ‘swineherd’
b. qul ‘kitchen gardener’
c. wak ‘street’
d. caj ‘tea’
e. qul ‘poultry farmer’
f. cajxana ‘printing plant’
g. caj ‘printing’
h. qul ‘prostitute’
i. wak ‘swineherd’
j. qul ‘kitchen gardener’
k. caj ‘tea’
l. qul ‘poultry farmer’
m. cajxana ‘printing plant’
n. caj ‘printing’
o. qul ‘prostitute’
p. wak ‘swineherd’
q. qul ‘kitchen gardener’
r. caj ‘tea’
s. qul ‘poultry farmer’
t. cajxana ‘printing plant’
u. caj ‘printing’
w. qul ‘prostitute’
x. wak ‘swineherd’
y. qul ‘kitchen gardener’
z. caj ‘tea’

<table>
<thead>
<tr>
<th>Chumash symbol</th>
<th>IPA</th>
<th>Phonetic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>[ts]</td>
<td>voiceless alveolar affricate</td>
</tr>
<tr>
<td>p'</td>
<td>[p']</td>
<td>ejective consonant</td>
</tr>
<tr>
<td>[ʔ]</td>
<td>[ʔ]</td>
<td>glottal stop</td>
</tr>
</tbody>
</table>

Do a complete morphological analysis of the data here.

i. For each lexical category (verb, noun, etc.), provide a list of roots in italics with their glosses in ‘single quotes.’

ii. For each category, provide a list of affixes (in italics, with appropriate hyphens) with glosses in ‘single quotes’ for translations, SMALL CAPS for grammatical terms.
Morphology: what's in a word?

- **r. ixut** 'be on fire'
- **s. kitswan** 'emerge'
- **t. saxkʰit** 'remember'
- **u. we’n** 'sleep'
- **v. susaxkʰit** 'remind'

Do a complete morphological analysis of the data here.

i. For each lexical category (verb, noun, etc.), provide a list of roots in *italics* with their glosses in *single quotes*.

ii. For each category, provide a list of affixes (again in *italics*, with appropriate hyphens) with glosses in *single quotes* for translations, SMALL CAPS for grammatical terms. If there is any allomorphy (changes in the shapes of any morphemes), first explain in words where each allomorph occurs, and then write a rule to describe the situation formally.

6. Burushaski is a language isolate (a language with no known relatives) spoken in Pakistan and India. (Data here come from Lorimer 1935, i:29–34.)

**SIDEBAR 4.21**

**TRANSCRIPTION NOTE**

<table>
<thead>
<tr>
<th>Burushaski symbol</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>[ʃ]</td>
</tr>
<tr>
<td>č</td>
<td>[tʃ]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ‘ruler’</td>
<td>thum</td>
</tr>
<tr>
<td>b. ‘crest of spur’</td>
<td>iriš</td>
</tr>
<tr>
<td>c. ‘millstone’</td>
<td>sil</td>
</tr>
<tr>
<td>d. ‘lamb’</td>
<td>ačas</td>
</tr>
<tr>
<td>e. ‘rope’</td>
<td>gusık</td>
</tr>
<tr>
<td>f. ‘bird’</td>
<td>bulis</td>
</tr>
<tr>
<td>g. ‘willow shoot’</td>
<td>yusık</td>
</tr>
<tr>
<td>h. ‘boy’</td>
<td>hikło</td>
</tr>
<tr>
<td>i. ‘hare’</td>
<td>saro</td>
</tr>
<tr>
<td>j. ‘butterfly’</td>
<td>holisło</td>
</tr>
<tr>
<td>k. ‘ox, bull’</td>
<td>hro</td>
</tr>
</tbody>
</table>

Do a complete morphological analysis of the data here.

i. Make a list of all roots in each lexical category (nouns, verbs, etc.). Give the roots in *italics*. Give glosses in ‘single quotes.’

ii. Make lists of all affixes that occur with each lexical category. Give the affixes in *italics*, with hyphens where appropriate. Give glosses in ‘single quotes’ if they are translations, and in SMALL CAPS if grammatical terms.

iii. If there are any changes in the forms of any morphemes when they are combined, describe in words the patterns you see. Write a rule or rules describing the changes.