What is Morphology?

Mark Aronoff and
Kirsten Fudeman

1 Thinking about Morphology and Morphological Analysis

mor·phol·o·gy: a study of the structure or form of something
Merriam-Webster Unabridged

1.1 What is Morphology?

The term **morphology** is generally attributed to the German poet, novelist, playwright, and philosopher Johann Wolfgang von Goethe (1749–1832), who coined it early in the nineteenth century in a biological context. Its etymology is Greek: *morph-* means 'shape, form', and *morphology* is the study of form or forms. In biology *morphology* refers to the study of the form and structure of organisms, and in geology it refers to the study of the configuration and evolution of land forms. In linguistics *morphology* refers to the mental system involved in word formation or to the branch

of linguistics that deals with words, their internal structure, and how they are formed.

1.2 Morphemes

A major way in which morphologists investigate words, their internal structure, and how they are formed is through the identification and study of **morphemes**, often defined as the smallest linguistic pieces with a grammatical function. This definition is not meant to include all morphemes, but it is the usual one and a good starting point. A morpheme may consist of a word, such as *hand*, or a meaningful piece of a word, such as the *-ed* of *looked*, that cannot be divided into smaller meaningful parts. Another way in which morphemes have been defined is as a pairing between sound and meaning. We have purposely chosen not to use this definition. Some morphemes have no concrete form or no continuous form, as we will see, and some do not have meanings in the conventional sense of the term.

You may also run across the term **morph**. The term 'morph' is sometimes used to refer specifically to the phonological realization of a morpheme. For example, the English past tense morpheme that we spell *-ed* has various morphs. It is realized as [t] after the voiceless [p] of *jump* (cf. *jumped*), as [d] after the voiced [l] of *repel* (cf. *repelled*), and as [əd] after the voiceless [t] of *root* or the voiced [d] of *wed* (cf. *rooted* and *wedded*). We can also call these morphs **allomorphs** or **variants**. The appearance of one morph over another in this case is determined by voicing and the place of articulation of the final consonant of the verb stem.

Now consider the word *reconsideration*. We can break it into three morphemes: *re-*, *consider*, and *-ation*. *Consider* is called the **stem**. A stem is a base morpheme to which another morphological piece is attached. The stem can be **simple**, made up of only one part, or **complex**, itself made up of more than one piece. Here it is best to consider *consider* a simple stem. Although it consists historically of more than one part, most present-day speakers would treat it as an unanalyzable form. We could also call *consider* the root. A **root** is like a stem in constituting the core of the word to which other pieces attach, but the term refers only to morphologically simple units. For example, *disagree* is the stem of *disagreement*, because it is the base to which *-ment* attaches, but *agree* is the root. Taking *disagree* now, *agree* is both the stem to which *dis-* attaches and the root of the entire word.

Returning now to *reconsideration*, *re-* and *-ation* are both **affixes**, which means that they are attached to the stem. Affixes like *re-* that go before the stem are **prefixes**, and those like *-ation* that go after are **suffixes**.

Some readers may wonder why we have not broken *-ation* down further into two pieces, *-ate* and *-ion*, which function independently elsewhere. In this particular word they do not do so (cf. **reconsiderate*), and hence we treat *-ation* as a single morpheme.

It is important to take very seriously the idea that the grammatical function of a morpheme, which may include its meaning, must be constant. Consider the English words *lovely* and *quickly*. They both end with the suffix *-ly*. But is it the same in both words? No – when we add *-ly* to the adjective *quick*, we create an adverb that describes how fast someone does something. But when we add *-ly* to the noun *love*, we create an adjective. What on the surface appears to be a single morpheme turns out to be two. One attaches to adjectives and creates adverbs; the other attaches to nouns and creates adjectives.

There are two other sorts of affixes that you will encounter, **infixes** and **circumfixes**. Both are classic challenges to the notion of morpheme. Infixes are segmental strings that do not attach to the front or back of a word, but rather somewhere in the middle. The Tagalog infix *-um-* is illustrated below (McCarthy and Prince 1993: 101–5; French 1988). It creates an agent from a verb stem and appears before the first vowel of the word:

(1)

Root	-um-	
/sulat/	/s-um-ulat/	'one who wrote'
/gradwet/	/gr-um-adwet/	'one who graduated'

The existence of infixes challenges the traditional notion of a morpheme as an indivisible unit. We want to call the stem *sulat* 'write' a morpheme, and yet the infix *-um-* breaks it up. Yet this seems to be a property of *-um-* rather than one of *sulat*. Our definition of morphemes as the smallest linguistic pieces with a grammatical function survives this challenge.

Circumfixes are affixes that come in two parts. One attaches to the front of the word, and the other to the back. Circumfixes are controversial because it is possible to analyze them as consisting of a prefix and a suffix that apply to a stem simultaneously. One example is Indonesian *ke . . . -an*. It applies to the stem *besar* 'big' to form a noun *ke-besar-an* meaning 'bigness, greatness' (MacDonald 1976: 63; Beard 1998: 62). Like infixes, the existence of circumfixes challenges the traditional notion of morpheme (but not the definition used here) because they involve discontinuity.

We will not go any more deeply here into classical problems with morphemes, but the reader who would like to know more might consult Anderson (1992: 51–6).

1.3 Morphology in Action

We would like to explore the idea of morphology more deeply by examining some data. These are examples of morphology in action – morphological facts of everyday life.

1.3.1 Novel words and word play

If you had been walking down the street in Ithaca, New York, a few years ago, you might have looked up and seen a sign for the music store "Rebop," a name that owes its inspiration to the jazz term *rebop*.[1] *Rebop* was originally one of the many nonsense expressions that jazz musicians threw into their vocal improvisations, starting in the early 1920s. In the 1940s, *rebop* became interchangeable with *bebop*, a term of similar origin, as the term for the rhythmically and harmonically eccentric music played by young black musicians. By the 1950s the name of this musical style was quite firmly established as simply *bop*.[2] Today, the original use of *rebop* is known only to cognoscenti, so that most people who pass by the store will be likely to interpret the word as composed of the word *bop* and the prefix *re-*, which means approximately 'again'. This prefix can attach only to verbs, so we must interpret *bop* as a verb here. *Rebop* must therefore mean 'bop again', if it means anything at all. And this music store, appropriately, specialized in selling used CDs. There's something going on here with English morphology. Of course, *rebop* is not a perfectly well-formed English word. The verb *bop* means something like 'bounce', but the prefix *re-* normally attaches only to a verb whose meaning denotes an accomplishment. The verb *rebop* therefore makes little sense. But names of stores and products are designed to catch the consumer's attention, not necessarily to make sense, and this one does so by exploiting people's knowledge of English in a fairly complex way and breaking the rules so as to attract attention, as verbal art often does.

Consider now the following phrases, taken from a Toni Braxton song: *Unbreak my heart, uncry these tears*.

We have never seen anyone *unbreak* something, and you certainly can't *uncry* tears, but every English speaker can understand these words. We all know what it means to unbreak somebody's heart or to wish that one's heart were unbroken. If we asked somebody, "unbreak my heart," we would be asking them to reverse the process of having our heart

broken. We can visualize "uncry these tears," too – we just think of a film running backwards. We can understand these words because we know the meaning of *un-*, which basically reverses or undoes an action. The fact that these particular actions, breaking a heart and crying tears, cannot be reversed only adds poignancy to the song.

All human beings have this capacity for generating and understanding novel words. Sometimes someone will create an entirely new word, as J. R. R. Tolkien did when he coined the now-familiar term *hobbit* (which, despite its popularity, is still not listed in the 2000 edition of the *American Heritage Dictionary*). But more often than not, we build new words from pre-existing pieces, as with *unbreak* and *uncry*. We could easily go on to create more words on this pattern.

Novel words are all around us. Jerry Seinfeld has talked about the *shushers*, the *shushees*, and the *unshushables* in a movie theater. Morley Safer was dubbed *quirkologist* – expert on quirky people – on a special episode of *60 Minutes*. For those who hate buffets, the TV character Frasier Crane came up with the term *smorgsaphobia*. Finally, the longest novel morphologically complex word we have been able to find on our own in the daily press is *deinstitutionalization*, from the *New York Times*.

These are everyday morphological facts, the kind you run across every day as a literate speaker of English. What these words – *rebop*, *unbreak*, *uncry*, *hobbit*, *quirkologist*, *smorgsaphobia*, and *deinstitutionalization* – have in common is their newness. When we see or hear them, they leap out at us, for the simple reason that we have probably never seen or heard them before. It is interesting that novel words do this to us, because novel sentences do not. When you hear a new sentence, you generally don't realize that this is the first time that you've heard it. And you don't say to yourself, "What a remarkable sentence," unless it happens to be one from Proust or Joyce or some other verbal artist. Many people have made the observation before that morphology differs from syntax in this way. **[Exercises 1–3]**

Morphological challenge

As you work through this book, keep an eye – or an ear – out for novel or otherwise striking words, on television, in magazines and newspapers, in books, and in conversations. Keep a running list of them, then e-mail your list to the authors: mark.aronoff@stonybrook.edu.

1.3.2 Abstract morphological facts

Now let's move to some more abstract morphological facts. These are the kind of morphological facts that you don't notice every day. They are so embedded in your language that you don't even think about them. They are more common than the ones we have just looked at, but at the same time deeper and more complex.

If you speak English and are concerned about your health, you might say:

(2) I eat one melon a day.

Let's imagine that we are even more concerned about our health than you are. We don't just eat one melon a day, rather:

(3) We eat two melons a day.

It is a fact about standard American or British English that we cannot say:

(4) *We eat two melon a day.

However, if we were speaking Indonesian or Japanese, we would say the equivalent of *two melon* (*three melon*, *four melon*, etc.) because these languages don't use morphological plurals in sentences like this.

(5) Indonesian:
Saiga makan dua buah semangka (se) tiap hari
I eat two fruit melon every day
'I eat two melons every day.'

Japanese:
mainichi futatsu-no meron-o tabemasu
every.day two- GEN melon-OBJ eat.IMPERF
'I eat two melons every day.'

The morphological grammar of English tells us that we have to put an *-s* on *melon* whenever we are talking about more than one. This fact of English is so transparent that native speakers don't notice it. If we happen to be speakers of a language without obligatory plural marking, however, we will notice it because we are going to have a lot of trouble with it.

We have now observed something about English morphology. If a word is plural, it takes the suffix *-s*. Living creatures don't eat only melons, however:

(6) The evil giant at the top of the beanstalk eats two melons, three fish, and four children a day.

Everyone agrees that *fish* is plural, but there is no plural marker. *Children* is also plural, but it has a very unusual plural suffix, *-ren*, plus an internal change: we say [ʧɪld-] instead of [ʧajld]. In other words, it's not always the case that we mark plural words with an *s*-like thing; there are other ways in which we can mark plurals. Native speakers of English know this, and they do not need to think about it before making a plural. **[Exercise 4]**

Consider the following:

(7) Today they **claim** that they will fix the clock tower by Friday, but yesterday they **claimed** that it would take at least a month.

In this example, we use two different forms of the verb *claim*. One is present tense, and the other is past. Again, this is not true for all languages. If we were speaking Vietnamese, for example, we wouldn't make any distinction between *claim* and *claimed* – we wouldn't mark the verb at all. If we were speaking Chinese, we would not distinguish between *claim* and *claimed* in a sentence like this, because the adverb *zuótiān* 'yesterday' is sufficient to indicate past tense:

(8) jīntiān tāmen shuō tāmen xīngqī wǔ ké yǐ xiū hǎo zhōnglóu,
today they say they Friday can fix well clock.tower
kě shì zuótiān tāmen què shuō zhì shǎo xū yào yíge yuè
but yesterday they however say at least need a month
'Today they claim that they will fix the clock tower by Friday, but yesterday they claimed that it would take at least a month.'

If we were to leave out *zuótiān* 'yesterday', we would need to use the particle *le* after the verb to show that the action took place in the past. In other words, whether or not a speaker must indicate past tense in Chinese depends on context.

Notice what happens in English when we use some other verbs besides *claim*:

(9) Today they **say** . . . but yesterday they **said** . . .

tell us	**told** us
know	**knew**

That these verbs and others do not add *-t*, *-d*, or *-əd* to make their past tense is an elementary fact about English morphology. We'll talk more about verbs like these later in the chapter.

The next observation about English morphology has to do with pronouns. The following is a real exchange between an American mother and her 6-year-old son:

(10) Who just threw a pool ball through the basement window?
Not me.

In this context, the 6-year-old (who was indeed guilty) would never have responded *Not I*. But if he were to answer with a sentence, the response would be *I didn't*, not *Me didn't*. In that case, the object form of the pronoun would be ungrammatical. Without formally knowing anything at all about subjects and objects, English-speaking 6-year-olds (and children even younger) master the pronoun system of their language. **[Exercise 5]**

Given the following sentence, how many children does Joan have?

(11) All of Joan's children are brilliant and play musical instruments surpassingly well.

From this statement you cannot know how many children Joan has, but one thing is certain: she has more than two. If Joan had only two children, we would normally say *both of Joan's children*, because it is a fact about English that there is a morphological distinction among universal quantifiers between the one designating all of two (*both*) or all of more than two (*all*) of a particular type of entity. In some other languages, marking for dual is even more pervasive. This is the case in Ancient Greek, as shown by the following examples:

(12) ho stratiô:tes lambánei tous híppous
the.NOM.SG soldier.NOM.SG take.3SG the.ACC.PL horses.ACC.PL
'The soldier takes the horses.'

to: stratió:ta lambáneton tous híppous
the.NOM.DU soldier.NOM.DU take.DU the.ACC.PL horses.ACC.PL
'The two soldiers take the horses.'

hoi stratiô:tai lambánousi tous híppous
the.M.PL soldier.PL take.3PL the.ACC.PL horses.ACC.PL
'The soldiers (three or more) take the horses.'

While English does not have special affixes to mark the dual, it keeps track of the distinction through words like *all* and *both*. There are actually languages in the world like Manam (Papua New Guinea: Gregersen 1976) and Larike (Central Maluku, Indonesia: Laidig and Laidig 1990) that distinguish not only singular, dual, and plural, but trial as well. The use

of singular, dual, trial, and plural second person subject prefixes in Larike is illustrated below:

(13) **Ai-** rala iter- lawa peʔa- o ?
2SG.SUB- chop.down 1PL.INCL.SUB- garden finish- QM
'Did you (sg.) finish clearing our garden?'

Kalu au- ʔanu, **irua** musti **iruai-** ʔanu siʔu.
if 1SG.SUB- eat 2DU certainly 2DU.SUB- eat also
'If I eat, certainly you both will eat too.'

Kalu **iridu-** ta- ʔeu, au- na- wela.
if 2TRI.SUB- NEG- go 1SG.SUB-IRR- go.home
'If you three don't want to go, I'm going home.'

Memang iri- hise tapi **imi-** ta- ʔariʔi-
truly 3PL.NONHUM- exist but 2PL.SUB-NEG- see-
ri.
3PL.NONHUM.OBJ
'They really do exist, but you (plural) didn't see them.'

1.4 Background and Beliefs

This book is a general introduction to morphology and morphological analysis from the point of view of a morphologist. The purpose is not to advocate any particular theory or to give the truth (whatever that is), but rather to get you, the reader, to where you can look for it by yourself. Still, it is inevitable that some of our remarks will be colored by our own beliefs and background. We would therefore like to present some of our foundational beliefs about linguistics and linguistic methodology.

First, we believe that **languages differ from one another**. You might be thinking, "Of course they do!" But we mean this in a very special way. Some linguists are always looking for ways that languages are similar, and at times, we do that, too. But we believe that if you focus only on the similarities between languages, you miss out on all of the exciting ways in which they differ. What's more, you may find parallels and similarities where none really exist. We try to approach linguistic analysis with as open a mind as possible, and to do this, it is first necessary to appreciate the uniqueness and diversity of the world's languages.

Our second foundational belief is that ***languages*, which we can write with a small *l*, are different from *Language*, with a capital *L***. There are

thousands of individual languages in the world. But we may also speak of language in general to mean the general phenomenon of Language that encompasses all individual languages. This Language is related to Noam Chomsky's notion of Universal Grammar, which posits that languages are all alike in basic ways. There is an important distinction between these two uses of the word *language* and each is equally important to linguistics. Individual languages have features that are not characteristic of Language in general. For example, one feature of English is that its regular way of forming plural nouns is to add /z/. We would never claim, however, that this is universally true, or that it is a property of Language. To tie this belief in with the preceding one, we strongly believe that morphological theory and morphological analysis must be grounded in morphological description. If we want to appreciate what morphology really is, it's best to have some idea of what the morphology of individual languages is like. At the same time, we must have a reasonably well-thought-out general theory of the morphology of Language, so that we can compare our descriptions of individual languages within a wider context. In short, linguists need to pay equal attention to both small-l language and capital-L Language.

Our next belief is that **morphology is a distinct component of languages or grammars**. If you are not already familiar with some of the controversy surrounding morphology, this needs an explanation. The fact that some languages, such as Vietnamese, do not have morphologically complex words has led some people to conclude that morphology should not be a separate branch of linguistics. The reasoning is that linguistics is generally understood to deal with properties of all languages – more precisely, Language with a capital L. If there are languages that don't have morphology, then morphology is not a property of all languages, and morphological phenomena should be treated in **syntax** or **phonology**. We disagree. It has been shown elsewhere (e.g., Aronoff 1994) that there are aspects of morphology that cannot be attributed to syntax or phonology, or anything else.

One piece of evidence that morphology is separate from syntax, phonology, and other branches of linguistics is that words in some languages are grouped into largely arbitrary classes that determine their forms in different environments. Latin nouns fall into five distinct classes, called declensions, which have little or nothing to do with syntax or phonology, and certainly cannot be explained by either. They are purely morphological in their significance. The uniquely morphological nature of these classes is truly brought home by the fact that Latin nouns also fall into

syntactic agreement classes (usually called genders) and the two systems cross-cut one another: two nouns may belong to the same gender but to different declensions and vice versa. We'll examine cases like these in later chapters, but their mere existence in many languages shows that morphology must be given some independent status in linguistics. Of course, morphology, probably more than any other component of language, interacts with all the rest, but it still has properties of its own.

We also believe that **morphologies are systems**. This is a very old observation. Because of it, it is impossible to talk about isolated facts in a language – everything holds together. This belief together with the second one, above, are the reasons why we'll be looking carefully at the morphology of a particular language, Kujamaat Jóola, throughout this book. Considering the morphology of Kujamaat Jóola in close to its entirety will give us a valuable perspective that we would never gain if we only focused on isolated facts from several languages.

So far, we have given you our beliefs about the nature of language and morphology. We also have some that pertain to methodology. The first is that we should **take an attitude of skeptical realism**. Albert Einstein said that a physicist must be both a realist and a nominalist, a realist in the sense that you must believe that what you ultimately find will be real, but a nominalist in the sense that you must never believe that you've found what you're looking for. Martin Joos made a similar statement about linguistics. On the one hand, you should always believe that what you are looking for is God's truth, but on the other, you should consider all that you have found so far as hocus-pocus. We believe strongly in the value of having a linguistic theory, but we believe equally strongly that you should never trust it completely.

Our other methodological belief can be summed up as a motto: **Anything goes**. This methodological belief is associated with the Against Method of Paul Feyerabend, a twentieth-century philosopher who felt that if we insisted on a single rule of scientific methodology, one that would not inhibit progress, it would be "Anything goes." We take a no-holds-barred approach to linguistics. We'll use any tool or method that will tell us how language works. This attitude stems in part from our skepticism about particular theories. People who are wedded to individual theories tend to believe in using tools that are rooted in that theory. Our tools are not theory-based in that way. If a tool does the job, we are happy to use it, whether it is a traditional linguistic tool (e.g., native speaker consultants, dictionaries, written grammars), an experimental tool (e.g., imaging technology), or a statistical tool.

1.5 Introduction to Morphological Analysis

1.5.1 Two basic approaches: analysis and synthesis

There are two complementary approaches to morphology, analytic and synthetic. The linguist needs both.

The analytic approach has to do with breaking words down, and it is usually associated with American structuralist linguistics of the first half of the twentieth century. There is a good reason for this. These linguists were often dealing with languages that they had never encountered before, and there were no written grammars of these languages to guide them. It was therefore crucial that they should have very explicit methods of linguistic analysis. No matter what language we're looking at, we need analytic methods that will be independent of the structures we are examining; preconceived notions might interfere with an objective, scientific analysis. This is especially true when dealing with unfamiliar languages.

The second approach to morphology is more often associated with theory than with methodology, perhaps unfairly. This is the synthetic approach. It basically says, "I have a lot of little pieces here. How do I put them together?" This question presupposes that you already know what the pieces are. So in a sense, analysis in some way must precede synthesis.

Say that you've broken a clock and taken it apart, and now you have to put all the little pieces back together. There's a catch: you don't know how. You could always go by trial and error. But the most efficient way would be to have some theory of how the clock goes together. Synthesis really involves theory construction.

From a morphological point of view, the synthetic question you ask is, "How does a speaker of a language produce a grammatically complex word when needed?" This question already assumes that you know what kinds of elementary pieces you are making the complex word out of. We think that one of the real problems of a morphological theory is that we don't always have a good idea of what the pieces are. Syntacticians can supply us with some tools: case and number, for example, are ancient syntactic notions that we can use in our morphology. But the primary way in which morphologists determine the pieces they are dealing with is by examination of language data. They must pull words apart carefully, taking great care to note where each piece came from to begin with.

We have described analysis and synthesis in terms of the morphologist studying language, but the two notions are equally applicable to speakers themselves. Speakers apply morphological analysis when they read or

hear a complex word that they have never encountered before. In order to understand it, they pull it apart and ask themselves whether they recognize any of the pieces. Speakers use synthesis whenever they create new forms from pre-existing pieces.

Read the caption in the following Motorola ad carefully. It contains an example of morphology in action – a striking morphological fact. Comment on it, relating it to the discussion of analytic and synthetic approaches to word-formation.

Who'd of thought that an electronic chip inside your car could help you avoid curbs, other cars, and best of all, Earl in repair.

1.5.2 Analytic principles

Before we encounter any actual problems, we would like to give you some basic analytic principles used in morphology. They are taken from Eugene Nida's (1949; revised edition 1965) textbook *Morphology*.[3]

The first principle is given in (14):

(14) **Principle 1**
Forms with the same meaning and the same sound shape in all their occurrences are instances of the same morpheme.

Step one in morphological analysis is to look for elements that have the same form and the same meaning. This is the basic type-token problem. Let's say that we have a bunch of coins. Each is a **token**, a form. If we look at them carefully, we see that three of them look very much the same (they are all nickels), and two of them are identical – they both say 1997. These two coins are tokens of exactly the same type: they have identical forms and identical values. We may further say that the three coins are all tokens of a larger type that includes all nickels, not just those minted in 1997. But five pennies, though they have the same value as a nickel, do not together comprise the same type as the nickel, because, although identical in value to the nickel, they are different in form.

Divide the following forms into morphemes. (For answers, turn the page.)

a. password
b. sprayable
c. childhoods
d. autobiography
e. co-educational

To apply this distinction between types and tokens to the morphological analysis of words, consider the Spanish words *buenísimo* 'very good' (< *bueno* 'good'), *riquísimo* 'very delicious' (< *rico* 'delicious'), and *utilísimo* 'very useful' (< *útil* 'useful'). In each case, the suffix *-ísimo* contributes the same superlative meaning, and it has the same shape. We logically conclude that the suffix is the same for all three words. Note that we presented three words, all with the same suffix. It is not enough to look at one form when attempting to break it up into its smaller parts. One thing that makes a morpheme a morpheme is that it recurs, and thus speakers are able to identify it and give it a meaning. **[Exercises 6–8]**

This isn't the whole story, as Principle 2 tells us:

(15) **Principle 2**
Forms with the same meaning but different sound shapes may be instances of the same morpheme if their distributions do not overlap.

In Kujamaat Jóola, for example, the stem /baj-/ has two possible shapes, [baj-] and [bəj-], but their distributions don't overlap. [bəj-] occurs in the presence of a morpheme with an underlyingly tense vowel, but [baj-] does not. This non-overlapping distribution allows us to conclude that the two forms are instances of the same morpheme. When two or more instances of a given morpheme occur with different shapes, we call them allomorphs. Allomorphs were introduced above in section 1.2.

The regular plural marker in English has several allomorphs – voiceless alveolar fricative /s/, voiced alveolar fricative /z/, schwa plus voiced alveolar fricative /z/, syllabic alveolar nasal /n/, and Ø – as shown in (16):

(16) seat-/s/
shade-/z/
hedg-/əz/
ox-/n̩/
fish-Ø

As in the previous example, the distributions of these forms do not overlap, and they all have the same meaning. We can infer that they are instances of the same morpheme.

(17) **Principle 3**
Not all morphemes are segmental.

Normally, when we think of morphemes, we think of forms that can be pronounced in some sense, e.g., *chicken*, *the*, *un-*, *-ize*. But some morphemes can't be pronounced on their own. They are dependent on other morphemes for their realization. In English, for example, vowel alternations may serve to differentiate basic and past forms of the verb. We refer to these alternations as **ablaut** (as in 18):

(18)

run	ran
speak	spoke
eat	ate

We know that there is a past tense marker distinguishing the words in the second column from those in the first. But what is it? It is not the /æ/

Answers to morpheme-breakup exercise:

a. pass/word
b. spray/able
c. child/hood/s
d. auto/bio/graph/y
e. co-/educ/at/ion/al

of *ran* or the /o/ of *spoke* but rather the difference between these vowels and the vowels of the basic verb, which is not segmental at all. We must look at both the present and past tense forms of these verbs, because it is the contrast between them that is important. Another type of non-segmental morpheme in English is shown in (19):

(19) breath_N breathe_V
cloth_N clothe_V
house_N house_V

In each pair, the noun ends in a voiceless fricative ([θ, s]), while the verb ends in a voiced fricative ([ð, z]). Assuming that the noun is basic, we say that the morpheme that marks the verbs consists of the phonological feature [+voice]. **[Exercise 9]**

Although Principle 3 says that we can apply the term morpheme to the non-segmental alternations seen in (18) and (19), it is nonetheless the case that doing so is awkward. Pairs like *run~ran* or *breath~breathe* are more easily explained as processes than as concatenation of morphemes. In the next chapter we will further address this issue. In section 1.2 we briefly mentioned classical problems with morphemes in the context of infixation and circumfixation. The existence of non-segmental alternations such as those in (18) and (19) is another classical problem.

The contrast between forms was crucial in (18) and (19). The notion of contrast can be further extended, leading to Principle 4:

(20) **Principle 4**
A morpheme may have zero as one of its allomorphs provided it has a non-zero allomorph.

Fish is an example of a word with a zero plural: one fish, two fish-Ø. We can say that it has a zero plural, and that this zero plural is an allomorph of the usual plural [z], because other words in the language, like *frogs*, have non-zero plurals. This is an analytic procedure, not a theoretical point. We cannot posit a zero unless it contrasts with some

non-zero variant. In Japanese, where *sakana* means both 'fish (sg)' and 'fish (pl)', we cannot posit a zero plural (**sakana*-Ø) because nowhere in the language does -$Ø_{PL}$ contrast with a non-zero allomorph. **[Exercises 10–11]**

1.5.3 Sample problems with solutions

Now that you have been introduced to some principles of morphological analysis, let us examine a data set. This one comes from the Veracruz dialect of Aztec, spoken in Mexico, and is taken from Nida (1965: 11):

(21) Aztec

a.	ikalwewe	'his big house'	i.	petatci·n	'little mat'
b.	ikalsosol	'his old house'	j.	ikalmeh	'his houses'
c.	ikalci·n	'his little house'	k.	komitmeh	'cooking-pots'
d.	komitwewe	'big cooking-pot'	l.	petatmeh	'mats'
e.	komitsosol	'old cooking-pot'	m.	ko·yameci·n	'little pig'
f.	komitci·n	'little cooking-pot'	n.	ko·yamewewe	'big male pig'
g.	petatwewe	'big mat'	o.	ko·yameilama	'big female pig'
h.	petatsosol	'old mat'	p.	ko·yamemeh	'pigs'

Our task is to list all the morphemes and to give the meaning of each. Before reading the following discussion, try this out on your own. Then, if you run into trouble or want to check your answers, read on.

We begin by looking for recurring pieces that have a consistent meaning or function. In this, the English glosses of each form are very useful. Consider (21a–c, j). All have something to do with 'house', and more specifically, 'his house(s)'. Examining the forms carefully we find that they all contain the piece *ikal-*, but have nothing else in common. We deduce from this that *ikal-* means 'his house'. We include a hyphen after *ikal-* because since it never appears on its own, we cannot know if Aztec requires that it be suffixed. The data set does not contain any other examples with an English gloss of 'his' or another possessive pronoun; nor does it contain any examples meaning 'house' without the possessor 'his'. This means that we cannot break *ikal-* up further.

Form (21a) *ikalwewe* 'his big house' contains the additional piece *-wewe*. Looking over the rest of the data, we find that *-wewe* also occurs in (21g) *petatwewe* 'big mat', (21d) *komitwewe* 'big cooking-pot', and (21n) *ko·yamewewe* 'big male pig'. All of these also contain the meaning 'big'. We conclude that *-wewe* means 'big'. Again we use the hyphen because in this particular data set, *-wewe* always appears attached to the stem.

One form contains the meaning 'big' but not the morpheme *-wewe*. This is (21o) *ko·yameilama* 'big female pig'. We recognize the piece *ko·yame-* 'pig', which also appears in (21m–n, p). Based on the minimal data we have, we can only deduce that like *-wewe, -ilama* means 'big', but that it attaches only to a certain class of noun. Both (21o) *ko·yameilama* 'big female pig' and (21n) *ko·yamewewe* 'big male pig' appear to have the same stem, but since one refers to a female animal and the other to a male animal, such a situation would not be unprecedented.

Based on (21b–c) *ikalsosol* 'his old house' and *ikalci·n* 'his little house' we isolate the pieces *-sosol* 'old' and *-ci·n* 'little'. This analysis is affirmed when we look at other words in the data set, such as (21e–f) *komitsosol* 'old cooking-pot' and *komitci·n* 'little cooking-pot', which contain the same pieces. We can also isolate *komit-* 'cooking-pot'.

In all, we can isolate the following morphemes:

(22)

ikal-	'his house'
komit-	'cooking-pot'
ko·yame-	'pig'
petat-	'mat'
-ci·n	'little'
-sosol	'old'
-ilama	'big' (occurs with stem meaning 'female pig')
-wewe	'big' (occurs with stems meaning 'his house', 'cooking-pot', 'mat', 'male pig')
-meh	plural marker

This exercise was fairly simple in the sense that there were no allomorphs, and the morphology was entirely morphemic – it did not interact with any non-segmental phenomena. However, there were a few difficulties. One was the fact that we did not have enough data to break up *ikal-* 'his house', and yet, since the English gloss clearly has two parts, you may have been tempted to break it into two parts, too. A second difficulty was the presence of both *-ilama* and *-wewe* 'big'. Again, we did not have enough data to understand fully what their difference is. Occasionally uncertainty is something that morphologists have to accept when working with published data sets and written grammars. Sometimes there are gaps in what is presented. Morphologists doing field research have the advantage of native speaker consultants whom they can ask. But in order to ask the right questions, it is important that morphologists alternate data collection with data analysis and not wait to get back home to analyze their findings.

A final observation is that this data set was not presented in the IPA. For example, you probably were not familiar with Nida's convention for marking long vowels: a raised dot, as in *-ci·n* 'little'. This fact in itself should not have posed any problems. It is often possible to isolate morphemes, particularly when there are no allomorphs or phonological interactions between them, whether or not we fully understand the transcription system. That was the case here. However, the use of a non-standard transcription system may make a problem set seem more daunting.

As explained in the prefatory remarks to this book, we chose to retain non-standard transcription systems despite the difficulties they present because as a linguist you will be faced with them time and time again. We hope that the experience you gain in this book will help you deal with such systems in your own research.

Our next sample problem set comes from French. It addresses different issues than the Aztec data discussed above. The French adjectives in the first column are masculine, and those in the second are feminine. Your task is to determine how masculine and feminine adjectives are differentiated and to outline a possible analysis. You may ignore changes in vowel quality.

(23)

	Masculine	Feminine	
a.	gros [gʁo]	grosse [gʁos]	'fat'
b.	mauvais [movɛ]	mauvaise [movɛz]	'bad'
c.	heureux [øʁø]	heureuse [øʁøz]	'happy'
d.	petit [pəti]	petite [pətit]	'small'
e.	grand [gʁɑ̃]	grande [gʁɑ̃d]	'big'
f.	froid [fʁwa]	froide [fʁwad]	'cold'
g.	soûl [su]	soûle [sul]	'drunk'
h.	bon [bõ]	bonne [bɔn]	'good'
i.	frais [fʁɛ]	fraîche [fʁɛʃ]	'fresh'
j.	long [lõ]	longue [lõg]	'long'
k.	premier [pʁœmje]	première [pʁœmjɛʁ]	'first'
l.	entier [ɑ̃tje]	entière [ɑ̃tjɛʁ]	'entire'
m.	gentil [ʒɑ̃ti]	gentille [ʒɑ̃tij]	'kind'
n.	net [nɛt]	nette [nɛt]	'clean'

As with the Aztec set, you should limit yourself to the data provided, although some of you may know French.

One way to begin is to see whether there is a single morpheme, which may or may not have allomorphs, that signals the difference between masculine and feminine. There is not. Masculine and feminine adjectives are differentiated by an alternation between Ø and [s] in (23a), Ø and [z]

in (23b–c), ∅ and [t] in (23d), ∅ and [d] in (23e–f), ∅ and [l] in (23g), ∅ and [n] in (23h), ∅ and [ʃ] in (23i), ∅ and [g] in (23j), ∅ and [ʁ] in (23k–l), and ∅ and [j] in (23m). The masculine and feminine forms of [nɛt] 'clean' are identical. (It is important to focus on pronunciation and not spelling. Spelling conventions are not part of the mental grammar.) We cannot consider the many final sounds of the feminine forms to be allomorphs of one another. Phonetically, they are extremely varied. Their distribution overlaps, too. For example, we find both [ʃ] and [z] after [ɛ], in (23i) *fraîche* 'fresh' (f) and (23b) *mauvaise* 'bad' (f), respectively. There is no apparent reason why (23n) *net*, *nette* 'clean' should behave differently from the other words in the list in having only one form [nɛt].

You may be thinking that the spelling can account for the final sound of the feminine forms. However, spelling often reflects the history of a word and not its synchronic analysis. Therefore, we cannot base our analysis on it.

So far it appears as if the final sound of the feminine forms of the adjectives is arbitrary. And yet, it cannot be wholly arbitrary, or speakers would not know which form the feminine takes. We have been treating this problem until now as if the feminine form is derived from the masculine one. A second possibility is that the opposite is true. We can form a hypothesis: perhaps the masculine form results when we remove the final sound of the feminine. This accounts for (23a–m). (Recall that we asked you to ignore changes in vowel quantity.) But this hypothesis fails when we apply it to (23n) *net*, *nette* 'clean'. Both are pronounced [nɛt]. Our current hypothesis, that we arrive at the masculine form by subtracting the last segment of the feminine form, cannot account for this fact.

At this point in the problem, you need to make a new hypothesis. There is room for more than one. One is that in French, adjectives have more than one stem, and both the masculine and feminine stems need to be memorized. This would mean that for (23b) *mauvais*, *mauvaise*, speakers memorize that the first is pronounced [movɛ] and the second [movɛz]. A second reasonable hypothesis is that we were on the right track earlier, and that speakers arrive at the masculine form by dropping the final segment of the feminine form. The feminine form is the only one that needs to be memorized, then, since the masculine can be derived from it by a regularly applying rule. Under this hypothesis, (23n) *net*, *nette* 'clean', both pronounced [nɛt], is an exception that speakers must memorize. Many would consider the fact that, based on our data set, this hypothesis requires speakers to memorize fewer forms to be an advantage.

We may not have arrived at a single, neat solution to the French data, but we have analyzed them and presented the hypotheses that they

suggest carefully. Presentation is important whenever you undertake to solve a linguistics problem. We close this section with a few tips for writing one up. First, when you include examples from a data set in the text of your analysis, set them off by underlining them or using italics, as we have done. Second, whenever you present a foreign-language form, provide its gloss, or definition. The most standard linguistic practice is to put the gloss in single quotation marks, like this: 'definition'. Finally, be sure you know what the problem is asking. If the problem asks for a list of morphemes, for example, that is all you need to provide (but don't forget to give their glosses, as well). If the problem asks for your analysis, present it carefully, as we have done above. In order to make your answer more compelling, you may need to explore analyses that do not work, as well. This is what we did in examining the French data. **[Exercises 12–14]**

1.6 Summary

We have given a whirlwind introduction to the field of morphology and to some of the phenomena that morphologists study. We introduced a key notion, that of the morpheme, but acknowledged that there are problems with its traditional formulation. We presented some basic beliefs of ours that underlie this and other chapters of the book, as well as four principles that will help the reader undertake morphological analysis. Finally, we led the reader through two sample problems in order to illustrate the steps a morphologist must take when analyzing data, as well as possible stumbling blocks that he or she might encounter.

We next turn to an introduction to Kujamaat Jóola, the language we have chosen to examine and analyze throughout this volume.

Introduction to Kujamaat Jóola

The Kujamaat Jóola people (who call themselves Kujamaat and their language Kujamutay) live in the Basse-Casamance region of Senegal, West Africa. Jóola is a cluster of dialects, of which Kujamaat, sometimes called Foñy, and Kasa are the most important.[4] The total number of speakers in 1998 was about 186,000 (Grimes 2002). Kujamaat Jóola belongs to the Atlantic (sometimes called West Atlantic) language family, of which the best-known languages are Wolof, the national language of Senegal, and Fula. Looked at in terms of linguistic history, the Atlantic languages form a branch descending from the most widespread language family in Africa, Niger-Congo, which is also one of the largest language families in the world. Kujamaat Jóola has a number of features – most particularly its intricate system of **noun classes** and agreement – which are remarkably similar to those of the distantly related but much larger and better-known subfamily of Niger-Congo, the Bantu languages.

The most pervasive and characteristic morphological features of Kujamaat Jóola are (i) a simple and elegant vowel harmony system, (ii) an extensive noun class or gender system, (iii) rich agreement morphology, and (iv) agglutinative verbal morphology. Over the course of this book we will be exploring these and other topics in Kujamaat Jóola morphology as they relate to issues raised in individual chapters.

We have chosen Kujamaat Jóola for this book because its morphology, though complex and sometimes unusual, is highly regular, which makes it an excellent teaching vehicle. The morphology is also spread out across nouns, verbs, and adjectives. The inflection includes some of the most common types that one is likely to find: nominal gender, agreement, and verbal tense and aspect. Finally, there is J. David Sapir's superb grammar, from which most of the Kujamaat Jóola data in this book are drawn, which provides a wonderfully lucid description of the language and especially of the morphology. The grammar has also stood the test of time: it speaks to us as clearly today as it did when it was written almost forty years ago.

Of all the distinct aspects of language, morphology is the most deeply entwined with the others. There is no way to talk about morphology without also talking about phonology, syntax, **semantics**, and **pragmatics**. Phonology is especially important, for there is no way to get at the morphology of a language without first stripping away the effects of phonology on the forms of words. For that reason our introduction to

Kujamaat Jóola morphology must be preceded by a brief overview of its phonology. Treatment of Kujamaat Jóola vowel harmony can be found in chapter 3.

The phonemic inventory of Kujamaat Jóola is given in (1) and (2).[5] Kujamaat Jóola has a set of voiceless and voiced stops in three places of articulation – bilabial, alveolar, and velar – and nasal consonants in four – bilabial, alveolar, palatal, and velar. It has voiceless and voiced postalveolar affricates /ʧ/ and /ʤ/, transcribed here as <c> and <j> (following Sapir 1965), voiceless labiodental and alveolar fricatives /f/ and /s/, two liquids /l/ and /r/, and labiovelar and palatal glides /w/ and /y/. The voiceless glottal fricative /h/ rarely occurs.

(1) Consonants

labial	alveolar	palatal	velar	glottal
p	t		k	
b	d		g	
m	n	ɲ	ŋ	
	c			
	j			
f	s			h
	l			
	r			
(w)		y	(w)	

Vowels occur in tense–lax pairs and may be short or long; what Sapir represents as schwa is realized as "a tense unrounded high-mid central vowel" under stress (Sapir 1965: 6), and is the tense counterpart to /a/. Tense high vowels are underscored (i̲ and u̲). The lax counterparts of tense /e/ and /o/ are /ɛ/ and /ɔ/ respectively:

(2) Vowels (all may be either long and short)

i	i̲		u	u̲
	e	ə	o	
		ɛ ɔ		
		a		

The organization of this vowel chart follows standard linguistic practice. It reflects the position of the tongue during articulation and resonance, with the high vowels [i, i̲, u, u̲] at the top of the triangle, and the low vowel [a] at the bottom. Vowels on the left are articulated toward the front of the vocal tract, and those on the right farther back.

Kujamaat Jóola words showing all of the vowels are listed in (3):

(3)

bəsi̲kən	‘mortar’
kəsi̲:t	‘feather’
gis	‘tear’
i:s	‘show’
ebe	‘cow’
-fe:gi̲r	‘three’
ɛfɛl	‘to untie’
ɛfɛ:l	‘to annoy’
ekəl	‘type of antelope’
ekə:l	‘to be partially ripe’
kafa:lɛn	‘to continue’
ɛgɔl	‘stick’
ɛgɔ:l	‘corner’
fu̲ko	‘head’
fu̲ko:k	‘wall’
ɛkuk	‘to take big handfuls’
ɛku:ku	‘mouse’
kəku̲ku̲l	‘to cultivate in dry ground’
kəku̲:ku̲:l	‘type of tree’

Nasal–nasal and nasal–consonant clusters are very common. Of these, only /mb/ and /nd/ occur freely, including at the beginning of a word; /nn/, /mf/ (transcribed here as <nf>, following Sapir), and /ns/ clusters occur only word-internally. The remaining clusters can occur in either internal or final position in a word. In all cases the two consonants have the same place of articulation. Both /lt/ and /rt/ occur in word-internal position, as well, though very rarely. There are no other consonant clusters. Some examples are given in (4):[6]

(4)

kəgu̲:mp	‘ashes’
mba	‘or’
nimammaŋ	‘I want’
-buntɛn	‘cause to lie’
nicɛɲcɛŋ	‘I asked’
-maɲj	‘know’
aŋkaŋk	‘hard’
emu̲ŋgu̲no	‘hyena’
fanfaŋ	‘lots’
ndaw	‘a man’s name’
-saltɛ	‘be dirty’
-ərti̲	‘negative suffix’

Kujamaat Jóola syllables are generally of the shape C(onsonant) V(owel), although V, VC, CVC, and CVNC (where N represents any nasal) syllables occur as well. Vowels may be long or short. Stress is stem-initial.

The most salient feature of Kujamaat Jóola phonology is its pervasive vowel harmony. Vowel harmony is the agreement among vowels in a word with respect to a given feature, such as height, rounding, or backness. We will explore Kujamaat Jóola vowel harmony in depth in chapter 3. Until then, keep an eye out for how certain morphemes influence the shape of Kujamaat Jóola stems, and, more often, vice versa.

Exercises

1. Create five new words – in English or your native language, if different. Give their definitions if they are not obvious.

2. Many product names are novel English forms coined by marketers. Look at the following list of product names and make hypotheses about how people came up with their names. Possibilities include, but are not limited to, the following: (i) combination of elements already occurring in English; (ii) combination of Latin or Greek morphemes – even without knowing Latin or Greek, you might be able to recognize a few; (iii) new use for a term already existing in English; (iv) use of a proper name. This is meant to be a fun exercise, ideally one to be discussed in class. It should not be graded.

   a. pHisoderm A pH-balanced cleanser
   b. Nescafé Coffee made by Nestlé
   c. Ajax A strong household cleanser
   d. Eucerin Moisturizing lotion
   e. Friskies Cat food
   f. Tums Antacid tablets
   g. Trident Chewing gum
   h. Life savers Hard candy shaped like a donut
   i. Spam Canned meat similar to ham

3. New technology creates a need for new words. You may not consider the TV remote control new, but relative to other examples of modern technology, it is. Besides "remote control," it is called by many other names. What do you call it? In class, compile a list of words that your classmates and instructor use to refer to it. Comment on the morphological form of the various words.

4. Choose a language other than English. It may be one you know or have studied, or one that you would like to learn more about by using library resources. How are nouns marked for plural in that language? Are they marked at all? Make a comprehensive list of plural types in the language, with examples.

5. We choose the example "Not me" and have it coming out of the mouth of a child (someone unlikely to have been exposed to much prescriptive grammar) on purpose. While some English speakers

may consider "Not I" to be more correct, many would agree that "Not me" sounds more natural. Can you think of other instances where "I" is considered to be more correct, at least by traditionalists, but where "me" sounds more natural, at least to you? What does this have to do with morphology, in your opinion?

6. Etymologically, the following words contain more than one morpheme. Break each of them up into its constituent morphemes, then list at least one other word that contains each morpheme. When identifying morphemes, it is always useful to identify other forms that contain them, and this exercise is to encourage you to begin doing so.

   Example:

morphemic	morph-	amorphous, polymorphic, metamorphic
	-em-	phoneme, hypoglycemia, academy
	-ic	tonic, sonic, academic

   a. monologue
   b. predispose
   c. receive
   d. phonology
   e. decline
   f. television
   g. circumscribe
   h. bibliophile

7. Rewrite the following forms and then separate them into morphemes using a slash or a hyphen. If a form consists of only one morpheme, call it **monomorphemic**.
   a. Danny
   b. theorists
   c. multifaceted
   d. weather
   e. monkey
   f. partnerships
   g. hysterical
   h. children

8. Should *-ful* be analyzed as one morpheme or two different morphemes (ful_1 and ful_2) in the following examples? Explain your answer and bring in further examples if necessary.
   a. wrathful
   b. handful

9. English noun and verb pairs
   A. The following words can be used as nouns or verbs, but their pronunciation changes accordingly. How? State your answer as

a generalization that contrasts the pronunciation of all the nouns with that of the verbs.

a.	import	g.	transfer
b.	contrast	h.	convict
c.	insult	i.	project
d.	insert	j.	rebel
e.	protest	k.	conflict
f.	convert		

B. For many English speakers, the verb *protest* has two different pronunciations. One fits the pattern that you identified in part A as being characteristic of nouns; the other fits the pattern you identified as being characteristic of verbs. If you are familiar with the two pronunciations of the verb *protest*, first identify the two possibilities, then come up with a hypothesis that might explain their coexistence.

10. English *spit* has two past tense forms: *spit* or *spat*. The second is an example of ablaut, mentioned in the discussion of Principle 3. What about the first? Should we analyze it as a single morpheme, or as two morphemes, *spit* and *Ø*?

11. Organize the following set of German nouns into singular–plural pairs. Then determine the allomorphs of the plural ending. Ignore changes in the stem vowel.

Väter	'fathers'	Auge	'eye'
Kinder	'children'	Adler	'eagle'
Pferd	'horse'	Kind	'child'
Männer	'men'	Augen	'eyes'
Vater	'father'	Kuh	'cow'
Mann	'man'	Frauen	'women'
Adler	'eagles'	Auto	'car'
Kühe	'cows'	Autos	'cars'
Pferde	'horses'	Frau	'woman'

12. Etymologically, the following forms contain more than one morpheme. In your opinion, does your mental grammar treat them as such, or does it treat them as monomorphemic forms? Deal with each form separately, because your answer may not be the same for all. Explain.
    a. holocaust
    b. parade
    c. presence

13. Zoque, Mexico (Nida 1965: 12)
List all morphemes and give the meaning of each.

pən	'man'
pənta?m	'men'
pənkəsi	'on a man'
pənkotoya	'for a man'
pənhi?ŋ	'with a man'
pənkəsita?m	'on men'
pənkəsiʃeh	'as on a man'
pənʃeh	'manlike'
pənʃehta?m	'like men'
nanah	'mother'
nanahta?m	'mothers'
nanahkotoya	'for a mother'
?unehi?ŋ	'with a child'
?unehi?ŋta?m	'with children'
naka	'skin, leather'
nakapit	'by means of leather'
nakapitʃeh	'as if by leather'
yomo	'woman'
yomota?m	'women'
yomohi?ŋ	'with a woman'
yomotih	'just a woman'
yomo?une	'girl'
kahʃi	'hen'
kahʃi?une	'chick'
libru	'book'
libru?une	'booklet'
wetu	'fox'
wetu?une	'fox whelp'
te? pən	'the man'
maŋu te? pən	'the man went'
maŋpa te? pən	'the man goes'
maŋke?tpa te? yomo	'the woman also goes'
minpa te? ?une	'the child comes'
minu te? ?une	'the child came'
maŋke?tu	'he also went'
maŋutih	'he went (and did nothing more)'

14. Congo Swahili, Elisabethville dialect
A. Identify as many morphemes as possible and give the meaning of each.

B. Imagine that you have the opportunity to do fieldwork on Congo Swahili. List a few sentences that you would elicit from consultants that might enable you to confirm or complete your morphological analysis.

Supplementary information:

a. The future *-taka-* and the negative *-ta-* are not related.
b. The final *-a* may be treated as a morpheme. Its meaning is not indicated in this set.
c. The passive morpheme may be described as having two forms, *-iw-* and *-w-*. Its form depends on what precedes it.

ninasema	'I speak'
wunasema	'you (sg) speak'
anasema	'he speaks'
munasema	'you (pl) speak'
wanasema	'they speak'
ninapika	'I hit'
ninanupika	'I hit you (pl)'
ninakupika	'I hit you (sg)'
ninawapika	'I hit them'
ananipika	'he hits me'
ananupika	'he hits you (pl)'
nilipika	'I have hit'
nilimupika	'I have hit him'
nitakanupika	'I will hit you (pl)'
nitakapikiwa	'I will be hit'
ninaona	'I see'
ninamupika	'I hit him'
tunasema	'we speak'
wutakapikiwa	'you (sg) will be hit'
ninapikiwa	'I am hit'
nilipikiwa	'I have been hit'
nilipikaka	'I hit (remote time)'
wunapikizwa	'you (sg) cause being hit'
wunanipikizwa	'you (sg) cause me to be hit'
wutakanipikizwa	'you (sg) will cause me to be hit'
sitanupika	'I do not hit you (pl)'
hatanupika	'he does not hit you (pl)'
hatutanupika	'we do not hit you (pl)'
hawatatupika	'they do not hit us'

NOTES

1 Conveniently, it also blended the first names of the two owners, Renee and Bob.
2 We thank Krin Gabbard for the etymology of *rebop*.
3 Nida has six principles; we present four here.
4 The Kujamaat Jóola data presented here comes almost exclusively from J. David Sapir's 1965 grammar, *A Grammar of Diola-Fogny*. We also used Sapir (1970, 1975; Thomas and Sapir 1967), Hopkins (1990), and Gero and Levinsohn (1993).
5 We choose to present the Kujamaat Jóola data in the transcription systems used by Sapir because being able to deal with different transcription systems is an essential skill for all linguists. Elsewhere in this book, we will generally use IPA transcription unless otherwise indicated.
6 In subsequent chapters nasal–consonant clusters will be written <nj>, <nc>, <ng>, and <nk>, respectively, following Sapir. In other words, we do not represent assimilation in place of the nasal to the following consonant (e.g., we write /nk/ for phonetic [ŋk]).