1

What Is a Word?

/ˈwʌt ɪz ə ˈwɜɹd/

In this chapter, we look at the intuitive notion of what a word is and see that there are several perspectives on wordhood. A word has different properties depending on whether you're looking at it phonologically, morphologically, syntactically or semantically. Essentially, we end up with two different notions of word: a *listeme* – a sound–meaning correspondence – and a *phonological word*, a sound unit on which the spacing conventions of written English are based. Finally, we distinguish between necessary and conventional aspects of wordhood.

1.1   Explaining *Word* in Words

Stop. Before reading any further, get out a sheet of paper and a pencil (or fire up a word processor, or just introspect), and try to compose a definition of the word *word*.

Exercise 1.1   Compose a definition of *word*.

Throughout this text, there will occasionally be exercises inserted in the middle of discussion. You should stop and try to answer them before reading on. Answers to the exercises are often given in the text immediately below; you'll be able to compare the response you came up with to the discussion in the text, and think about any differences between the answer in the text and your own answer.

1

What Is a Word?

Here's one possible first try:

Definition 1
word: a sequence of letters that we write consecutively, with no spaces.

How does that definition compare with your own? Yours is probably better. One thing that is obviously wrong with this one is that it depends crucially on the conventions of writing. Languages have words before they're written down. Let's try again, trying to eliminate the reference to writing:

Definition 2
word: a sequence of sounds that we pronounce consecutively, with no pauses.

Hang on a minute – when we're talking, there's not usually any pauses between words. (Try listening for a moment to someone talking. Is there a pause before and after every word? Where *are* the pauses?) We do know, though, that it is at least possible to put pauses between words when talking. Imagine you are speaking to someone for whom English is a second language, and who is hard of hearing besides. To give them the best chance of understanding you, you . . . would . . . probably . . . talk . . . rather . . . like . . . this, inserting big spaces between words. (People talk like this when dictating, as well.) You certainly wouldn't insert spaces inside them. No one would say "y . . . ou . . . wou . . . ld . . . pro . . . b . . . abl . . . y . . ." etc. Maybe we can use the *possibility* of spaces in our definition:

Definition 3
word: a sequence of sounds which *can* be pronounced on its own, with pauses on either side.

Hang on again! A word is not just any old sequence of sounds that can be pronounced on its own. According to that definition, *spimble* or *intafulation* or *pag* are words, and so are *raise your arm* or *how are you* (you can pronounce them with space on either side, can't you?). The former, however, are sequences of sounds that don't have any meaning associated with them, and the latter are sequences of sounds that have too much meaning associated with them. Intuitively, the former are not words, and the latter are groups of words.

2

To help make the text clearer, when we're discussing the linguistic properties of some word, the word will appear in italics. This indicates that the word is just being mentioned – that is, being discussed – rather than being actually used. This mention/use distinction is hard to keep track of when it's not indicated by some distinguishing feature, such as italics.

It seems fairly clear that we have to include meaning in our definition. The sounds that make up, for instance, the word *word* have a certain meaning in combination that they don't have by themselves, or when they appear in other words (like *water* or *murder*). So the *w* sound in *word* doesn't mean anything by itself, nor does the *-ord* sequence, but together, they have a meaning, even if it's a meaning that's hard to pin down. So for our final try, let's look at the relevant definition in the *Oxford English Dictionary* (**OED**), which is listed as definition number 12a in their entry for the word *word*:

Definition 4 (final)
word: A combination of vocal sounds, or one such sound, used in a language to express an idea (e.g. to denote a thing, attribute, or relation), and constituting an ultimate minimal element of speech having a meaning as such; a vocable.

This is probably fairly close to the definition you came up with, albeit perhaps with a few extra elements. The crucial part that we didn't have in our earlier versions is the bit about the "ultimate minimal unit of speech having a meaning as such."

So consider our example word, *word*. The *w* doesn't have a meaning by itself, nor does any other individual sound. The first three sounds, which we spell *wor* in the word *word*, do have a meaning of their own (spelled *were*, the past plural of the verb *to be*), but that meaning is not a part of the meaning of *word* – that is, the meaning of *word* does not include the meaning of *were*. Other subsets of the sound sequence (*or*, *rd*, *ord*) are similarly unrelated in meaning or meaningless. *Word*, then, is a minimal unit of speech having a meaning.

This definition works to eliminate our counterexamples above from consideration as possible "words." *Spimble*, *intafulation* and *pag* are units of speech that don't express any idea, and *raise your arm* and *how are you* are units of speech that have a meaning, but they aren't

3

minimal – their meaning is made up of the meanings of the smaller elements within them, each of which contributes its own meaning to the meaning of the whole expression in a consistent way. So although the meaning of *were* is not part of the meaning of *word*, the meaning of *raise* IS a part of the meaning of *raise your arm*.

Nonetheless, we'll see that this definition of *word* does not correspond with the everyday sense of the word *word* in English.

Exercise 1.2 Can you figure out why this definition doesn't match the usual meaning of "word" before reading Section 1.3? Try to think of English words or expressions which are counterexamples.

Before we do that, however, let's look at basic design of language, in order to understand the central role that words play every day in the dance of communication.

1.2 Language Is a Secret Decoder Ring

Language lets us see into other people's minds, and lets other people see into ours. If we speak the same language, then just by talking, I can cause you to have an idea that I have had, or at least a close approximation of it. If we speak different languages, no amount of talking will let me share my idea with you. It's as if learning a language is like getting a secret decoder ring that lets you encrypt thoughts and feelings and transmit them to someone with the same decoder ring. What's especially great about this encryption device that we all carry around in our heads is that it's more or less automatic. You don't (usually) have to *consciously* identify and match up the symbols (the spoken words) to the ideas; it happens automatically, both on the sending end and the receiving end.

Consider the stick figures modeling the communication process in Figure 1.1. The skirted figure has an idea to communicate (panel 1). She encodes it into a linguistic form (panel 2) – ultimately, a string of instructions transmitted by her nerves to her vocal cords, lips, and tongue – and creates some sound waves (panel 3). The stick figure she's talking to hears the sound waves (panel 3), translates them back into an abstract linguistic form (panel 4), and ultimately, back into the idea (panel 5).

Even though it doesn't take very long to accomplish the encryption in step 2 and the decryption in step 4 of this process, it's an incredibly
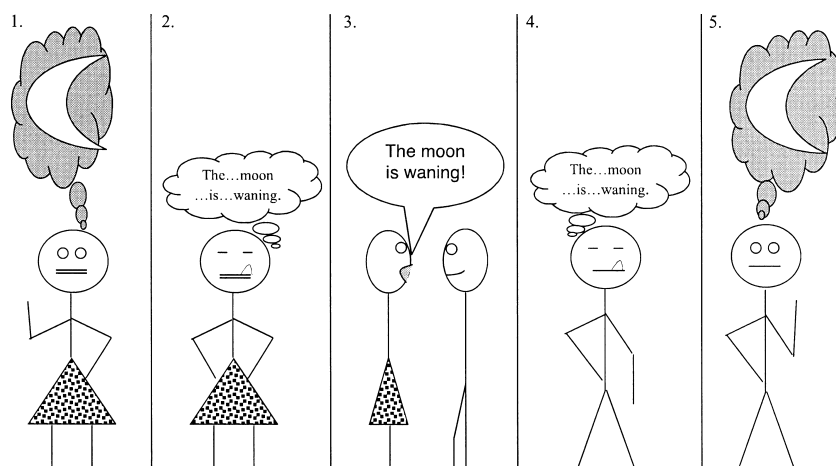
Figure 1.1 Communicating using language

complicated business. (This book is mostly about just one sub-part of what's involved during this process, the part that has to do with words.)

The encryption system has two basic parts. The first part is a set of symbols which stand for concepts, like the English word *dog* is a symbol standing for the concept DOG. (Note that in French, the word *chien* stands for the concept DOG, in Spanish, *perro* stands for the concept DOG, and in Hiaki, a language spoken in southern Arizona and northern Mexico, the word *chu'u* is the symbol that stands for DOG.) These symbols are, of course, words. In spoken language, words are made up of sounds produced by the vocal cords, lips and tongue, but they don't have to be: sign languages use certain handshapes and motions as the building blocks of words. Any symbol can behave like a word if it's associated with an appropriate meaning.

You can get pretty far, communication-wise, with just words, even without the second part of the encryption system. Chimpanzees trained in sign language can do pretty well at communicating ideas about their likes and dislikes, needs and wants, and about things in the immediate environment, using unstructured clusters of words. The second part of the encryption system, though, is what makes it infinitely versatile. There's a procedure for sticking symbols together to make up complex units that correspond to complex ideas: the meanings of the complex units derive from both the meanings of the symbols (part one) AND the rule used to combine them (part two). Crucially, these combining rules are *recursive*: they can construct complex units that contain other

5

complex units of the same type (*this is the cat that chased the rat that ate the malt that lay in the house that Jack built*). Because they are recursive, these rules can create infinitely long and complex sentences. The rules are called *syntax*. By combining meaningful symbols in a structured, hierarchical way, syntax allows us to communicate about our plans, our beliefs, our hopes and fears, and our procedure for replacing a timing belt in a 1999 Toyota pickup truck.

So the skirted figure in step 2 of Figure 1.1 is doing two things: (1) selecting the right words for the concepts that make up the sub-parts of our idea; and (2) selecting the right combination of rules to stick the words together so that they add up to the idea she's trying to get across. The syntactic rule system is what lets us encode and understand the differences between *a dog is barking* and *a dog that is barking* and *a barking dog* and *there is a barking dog* and *there is a dog that is barking* and *the dog that is barking is barking* and *a barking dog is barking* and *a barking dog that is barking is barking* . . . and so on.

Compare the following two strings of words:

(1)   The dog that is barking

(2)   The dog is barking.

The only difference between them, word-wise, is that the first group of words has one more word in it than the second. Nonetheless, they mean fundamentally different things to an English speaker: the second one is a complete sentence describing an event that is happening right now, while the first one is a phrase that refers to a particular being in the world – a noun phrase – but it is not a complete sentence.

Now compare these two strings of words:

(3)   *Is dog the barking that

(4)   *Is dog the barking.

Here and throughout this book we will use the asterisk symbol * in front of examples to indicate that they are ill-formed, or *ungrammatical* in the linguist's sense. (In this use, the symbol is called a "star.") Examples marked with an * sound funny. It's not that they are stylistically disfavored, like *ain't* or *Where did the cockroach run to?* They are simply not produced by the linguistic system of a speaker of English.

These two strings are made up of exactly the same words as the first two, and differ in exactly the same way, word-wise – (3) has one more word in it than (4). However, the extra word – "that" – has much less effect in these two strings of words than in the first two: both of them are just gibberish, with or without the "that." You can recognize that the individual words mean something, but it's hard to tell whether the whole string of words means anything at all, let alone whether (3) means something different from (4). This is the effect of the second part of the encryption system. It is the way the words are put together – their syntax – that makes the sequences in (1) and (2) so different from the sequences in (3) and (4).

We'll learn more about both parts of the system as we go along, and how the parts interact, but for now, let's get back to our central question for this chapter. What's the problem with defining a "word" as "the minimal unit of speech with its own meaning"?

1.3 Wordhood: The Whole Kit and Caboodle

1.3.1 *Minimal units with meaning that are smaller than "words"*

Here's the problem: there are many cases where an "ultimate minimal element of speech having a meaning" is *smaller* than the units we put spaces around when we're writing or talking slowly, i.e. the ultimate minimal unit of meaning can be smaller than the things we normally refer to as "words." Let's take a fairly straightforward case first. Read the sentences below aloud to yourself:

(5) a. I'm mad at you.
    b. Don't take candy from strangers.
    c. Why couldn't you carry it more carefully?
    d. You aren't going out dressed like that, are you?
    e. You're not going out dressed like that, are you?

Exercise 1.3 What is it about these sentences that poses a problem for defining "word" as an "ultimate minimal unit of speech having a meaning"?

7

In each of (5)a–e, it should be clear that there is an element that is surrounded by space on both sides (and that can be pronounced as a word on its own), but that single element contains two concepts – two units of meaning. That is, as pronounced (and written), they count as single words, but they are combinations of two elements as far as meaning is concerned. The items in question in (5)a–e, plus several other common examples, are listed in (6):

(6)   I'm, don't, couldn't, aren't, you're, he's, they've, we're . . .

Of course, you might argue, these aren't true counterexamples to the definition, because they are **contractions**, squeezed-together versions of two real words, both of which constitute minimal units of speech with meaning in their own right. *I'm* corresponds to *I am*, *don't* is *do not*, *you're* is *you are*, *aren't* is *are not*, etc. On some level, then, these are truly separate words, and this is reflected in that they *can* be pronounced as separate words. At some point during linguistic processing and before actual pronunciation (in panel 2 in Figure 1.1), the two words get pushed together and are pronounced as a single unit. In order to make the OED definition match up to our everyday sense of "word," then, it needs to be altered. What if we say that a "word" isn't *always* a sequence of sounds that is pronounced separately (an "ultimate minimal element of speech"), but rather, it's a phonological unit that *could* be pronounced as a separate sequence of sounds, as we did in our third definition revision above? Then in the sentences above, *n't*, *'re*, and *'m* would count as words, because they could have been pronounced *not*, *are* and *am* instead.

   If we make that move, we take care of another troublesome class of words: *compounds*, words made up of two words in combination. Some good examples are *homeowner*, *blackbird*, *man-eater*, *greenhouse*, *overhead*, *pickpocket*, etc.

   This revision isn't enough, however. Contractions and compounds are not the only ways that two meanings, attached to two sets of sounds, can be packaged up into a single word. Consider the word *dog*, which is a word that satisfies the definition: none of the possible minimal units contained within the word (*d*, *do*, *o*, *og*, *g*) have any meaning of their own (or no meaning that contributes to the meaning of the whole), so *dog* is a minimal unit of speech with its own meaning – it doesn't get any of its meaning from some smaller unit within it. Now, what about the word *dogs*? Its overall meaning appears to be

8

made up of two elements: the word *dog* that we just saw, plus a **suffix** *-s*. As a speaker of English, you will know that the *-s* suffix, applied to **nouns**, indicates plurality – it means, roughly, "more than one X," where X is the noun it's attached to. In the dictionary, it could have that as its definition like this:

-s: More than one X (where X is the noun *-s* is attached to)

So here we have a sound unit, *-s*, which has its own meaning, PLURAL, and yet it's certainly not anything that we would call a "word" on its own – it can't be pronounced by itself in answer to a question, for example:

(7) *Jack*: How many of them did you see?
     Jill: * S. (*intended meaning*, "More than one.")

Of course, any suffix with a regular meaning falls into this category. In (8) we see some groups of words with prefixes and suffixes, whose meanings are regular combinations of the meanings of their various parts:

(8) a. iconic, acrobatic, idealistic, photographic, idyllic, robotic
     b. writing, hammering, presenting, kissing, analyzing, shivering, thinking
     c. bendable, breakable, manageable, loveable, fixable
     d. unbeaten, unhappy, un-American, unwanted, undefined, unremarkable
     e. writer, gardener, clipper, timer, greeter, cleaner, washer, dryer

---

Exercise 1.4 Based on these lists of words, see if you can come up with a definition for each of the affixes *-ic*, *-ing*, *-able*, *un-* and *-er* shown in (8)a–e, on the model of the definition given above for *-s*.

So, there are minimal sound sequences that have meaning that cannot stand on their own. Such sound sequences are not words as we use the term in everyday language – we don't write them with spaces on either side, like this: *dog s*, *icon ic*, *bend able* – nor, if we are spacing "words" apart and speaking slowly, do we include pauses between the pieces.

9

> **phonology**, *n.* From the Greek roots *phono-*, "voice, sound" and *-logy* "saying, speaking." 1. The study of spoken sounds. 2. The system of sounds in a language. **phonological**, *adj.* relating to phonology.

A *phonological word* is sequence of sounds which is identified as a unit on the basis of how it is pronounced – a collection picked out by the phonology of a language. *Can't*, *bendable* and *dogs* are phonological words.

1.3.2 Phonological words that don't carry any meaning whatever

In addition to the problem posed by affixes, above, there's another problem for the definition we're considering, although examples are somewhat harder to come by. Consider the following phrases:

(9) a. Jill took it all, *kit and caboodle*
  b. Jack walked *to and fro*
  c. If I *had my druthers*, the party would be on Saturday.
  d. The responses *ran the gamut* from brilliant to insane.

While it's clear to most speakers of English what the phrases *kit and caboodle*, *to and fro*, *have (one's) druthers*, and *run the gamut* mean (respectively, "everything," "back and forth," "get one's way," and "varied as widely as possible"), hardly any speakers know what the words *caboodle*, *fro*, *druthers*, or *gamut* mean in these expressions (no one would ever say "Do you like John's druthers?" or "She made it clear she wanted the caboodle."). Perhaps a guess can be made about the meaning of *fro*, since the phrase is so much like *back and forth* in structure and meaning: it seems like it must mean the same thing as *forth*. Yet, *to and forth* is nonsensical, and *forth* in other uses cannot be replaced by *fro*. Who ever heard of a knight going fro on a quest? Yet, *fro, caboodle*, etc. clearly are phonological words, shown by the fact that they can be pronounced, and are written, with spaces on either side. Essentially, what these examples show is that there can be phonological words which don't have a meaning associated with them at all, but only acquire meaning in conjunction with other phonological words. According to

10

the OED definition, however, *kit and caboodle* is one "word," as it is a minimal unit of speech having a meaning. Do you agree?

It's not simply that there are some phonological words that have no meaning. There's an enormous class of expressions made up of several phonological words that do have meanings but whose meanings have nothing to do with the meaning of the whole expression. Consider the examples in 0:

(10)   a.   Bill *kicked the bucket* last night.
         b.   The promotion is a real *feather in her cap*.
         c.   Fred was suffering from an attack of *the green-eyed monster*.
         d.   He wouldn't stop complaining, but he was *flogging a dead horse*.

There's no actual, or even metaphorical, bucket involved in (a), no feather, monster or horse in (b), (c) and (d). These phrases are **idioms**, expressions whose meaning must be learned by rote, just as one would learn the meaning of *pith* or *reimburse*. As they occur within these expressions, these phonological words have no meaning associated with them at all: the only meaning around is associated with the larger phrase of which they form a part. Since these phrases are minimal units of meaning, but are composed of many smaller, easily identifiable phonological words – minimal units of speech – they too show that "word" cannot be defined as something that correlates a minimal unit of speech with a minimal unit of meaning.

1.4   Two Kinds of Words

There's an easy way out of this dilemma. On one view, the meaning of "word" has mainly to do with semantics – the part of the definition that refers to the "minimal meaningful unit," that is, an element of the list of sound–meaning correspondences that is one of the two fundamental elements of language. The other, more everyday interpretation of the meaning of "word" has mainly to do with **phonology**: the fact that we call whatever we can pronounce in isolation a "word." The latter we have simply labeled: *phonological word*. We'll learn some of the properties that English requires of its phonological words in Chapter 2. The former, the true *minimal meaningful unit*, which includes affixes, like *-s* and *un-*, and idioms like *kick the bucket*, we will call *listemes*.

11

Listemes are often also called *morphemes*. We'll learn more about morphemes soon, and discuss why in this volume we distinguish listemes from morphemes. Another technical word that also has a similar meaning is *lexeme*. It could be useful to look these terms up in several different linguistic encyclopedias, dictionaries, or glossaries and compare their definitions and uses.

Why "listemes"? Since these sound–meaning combinations are *arbitrary*, the connection must be *listed* in the speaker's (your) head somewhere. We know that listemes are arbitrary because languages use different words for the same concept (as we saw in the names *perro*, *dog*, *chien*, and *chu'u* for the concept DOG, above). Indeed, any group of people – say, a children's secret club – could just get together and decide: "We won't call this a *dog* anymore, it's now a *spimble*." Similarly, while it would be considerably more difficult to stick to, a secret club could equally decide that they wouldn't make plurals with *-s* anymore; rather, they'd use *-int*. ("Mom! Where's my box of colored pencilint?") Ferdinand de Saussure called this property *the arbitrariness of the sign* (Saussure, [1916] 1959). Another way of putting it is that there is no "right" name for any concept, except what speakers of a language happen to agree on. This list of items is what learners of second languages spend hours memorizing, and it's what dictionary makers try to replicate. (Look in any college or unabridged dictionary. It includes not only phonological words *per se*, but also many affixes and idioms: there'll be an entry for *-ed*, one for *un-*, one for *-ing*, etc.). This book is about phonological words and listemes, and their love–hate relationship.

1.5   The Anatomy of a Listeme

Stop again. Before reading any further, make a list of the *minimum* amount of information you think it is necessary to know in order to know the (most common meaning of the) word *nice* and use it like an English speaker. (No looking in the dictionary, now. What do *you* know about it? Imagine you had to explain this word to someone learning English so that they could use it in speech.)

12

Exercise 1.5 Make a list of the minimum amount of information it is necessary to know in order to "know" the word *nice*.

Here are some things that all English speakers know about *nice*:

1 **Pronunciation**. You know how to pronounce it. A set of instructions for pronouncing the word *nice* might go like this: First, press the tip of your tongue to the roof of your mouth behind the tongue, blocking off all air exiting through the mouth Create a sound by allowing air to escape through your nose while simultaneously tightening your vocal folds so that the air passing over them causes them to vibrate. Then, continuing to vibrate your vocal cords, open your mouth with your tongue almost flat, allowing air to escape. Raise your tongue up and forward somewhat, vibrating your vocal cords all along. Finally, bring your tongue nearly all the way to the top of your mouth behind the teeth, creating a narrow opening. Stop vibrating your vocal cords and allow air to pass through the opening, making a hissing noise as it does so. (Isn't it lucky we don't have to have this kind of instruction to learn to talk? In any case, it's clear that all of this is information you know about *nice*.)

2 **Meaning**. You know what it means: something like "pleasant, agreeable."

3 **Category**. You know that it is an **adjective**. That is to say, even if you've never heard the word *adjective*, you know that *nice* can modify nouns (*a nice picture*). *Adjective* is just a term that means roughly "a word that can modify a noun." Speakers of some dialects of English also use it as an adverb, to modify verbs (*he sings nice*), so if you speak such a dialect, you can list "adverb" next to "adjective" as something that you know about *nice*.

4 **Other forms**. You know that it consists of a single, stressed syllable, and hence that it has a **comparative** form *nicer*, and a **superlative** *nicest*. (This is not true of all adjectives: compare *nicer* to the comparative form of *aware*: *more aware*, not *awarer). If you speak a dialect like Standard American English that doesn't allow *nice* as an adverb, you can also list the adverbial form *nicely* as something you know about *nice*.

How much of the above was in your list? You might have spent the most time on 2, and you might have omitted to mention any of 1, 3

13

and 4 entirely. Nonetheless, anyone who speaks English and has the word *nice* in their vocabulary certainly knows all of the above. All of this information must be in your head somewhere.

In traditional linguistic study, the information in 1, about pronunciation, is part of *phonology*. In 2, the information about meaning is part of **semantics**. In 3, the information about category is part of *syntax*. And finally in 4, the information about affixes and the internal structure of the word is *morphology*. When a child (or anyone) learns a new listeme, they learn (or figure out) at least some information from all of the above categories. They *have* to; that's what it means to learn a word.

1.6   What *Don't* You Have to Learn When You're Learning a Word?

Many of you might know a great deal more about the word *nice*. For instance, I'm fairly sure that everyone reading this textbook knows how to *spell* the word nice. Stop and consider a moment, however. Is it necessary to know how to spell a word to "know" it? Consider a 5 year old, who can't read or write. After hearing his mother read *Jack and the Beanstalk*, he says, "That was a nice story." He certainly can't *spell* the word "nice," but would you say he doesn't know the word "nice"? It seems clear that he *does* know it, enough to pronounce it correctly and use it accurately.

Some of you might know something about the history of *nice*. It was borrowed by English speakers from Old French in about 1300 AD, and originally meant "stupid or foolish," which is what it meant at the time in Old French. Over the years, it went through many permutations of meaning: from "foolish" to "loose-mannered, wanton," and from there to "lazy, indolent, slothful." From "lazy" it permuted to "not able to endure much, delicate," and thence to "over-refined." Then it was a short step to meaning "fastidious, difficult to please," which became, "precise, finely discriminating," which became "refined," and, applied to food, "dainty, appetizing," which finally led to our modern sense, "agreeable, pleasant" (with several side-shoots of meaning that I haven't mentioned).

In Old French, *nice* had developed over the years from the Latin word *nescius*. *Nescius* in Latin was originally a contraction of the phrase *ne scire*, "not to know" (hence, "stupid, foolish"). The Latin verb *scire*,

14

meaning "to know," is also the root of the English word **sci**ence, as well as *pre**sci**ent*, *con**sci**entious*, *omni**sci**ent*, and *con**sci**ous*, although these were borrowed by English at a much later date than *nice* was.

Some of you might know that *nice*, while quite a nice word, is used so frequently that some sophisticated writers of English consciously try to avoid it: a sentence that is stylistically strong and descriptively gripping doesn't usually have the word *nice* in it. If you're a speaker of a dialect of English which allows *nice* as an adverb, as in *She sings nice*, you may also know that Standard English – the English you are expected to use in written work at school or in professional settings – does not permit *nice* to be used as an adverb.

The above information, while interesting and true, is not part of what anyone automatically learns when they're first learning the word *nice*. We'll be learning about both types of knowledge in this book: the complex information about words that all English speakers carry around in their heads, and the historical and social information about words that is the result of accidents of history and language change. The former information tells us about the nature of our minds, giving us a window onto the computation that goes into the utterance of the simplest English sentence; the latter information can give us an insight into the history and culture of the people who have spoken and written English over the past 10 centuries. We'll be talking about both kinds of information, but we'll be taking care not to get them mixed up. The first kind of information belongs to the study of psychology of language, and the latter to the study of the history of language. Keep the distinction in mind as we go on. If you're wondering which category a certain kind of information falls into, ask yourself: is this something that children who speak English know?

The study of the psychology of language and study of the history of language are connected by the study of the *sociology* of language, the study of how and why people end up speaking the way they do. Psycholinguistics, historical linguistics and sociolinguistics are all subdisciplines of linguistics, areas in which a linguist can choose to specialize.

1.7   A Scientific Approach to Language

In this book, we will be studying English words in the same way an entomologist would study a species of insect, the same way a geologist would study layers of rock, the same way a meteorologist would study weather patterns. We will look at English, describe what we see, and then try to develop an analysis that explains any patterns or regularities that we find.

We won't be concerned, in our study, with "correct" or "proper" ways of speaking and writing English, except insofar as they are relevant to our discussion of how people actually do speak or write. Teaching English speakers to adhere to certain rules of grammar, or punctuation, or style, is undertaken by people interested in a *prescriptive* approach to English, who are interested in ensuring conformity among speakers of English for some purpose. We here taking a *descriptive* approach: trying to discuss what English speakers actually do, not what they "should" do.

If you are a second language learner of English, this book will be useful in your study of English: it is full of information about what native English speakers actually do when they're speaking English. If you are a native speaker of English, you will find that this book tells you about *how* you speak English, and something about *why* modern English is the way it is – but it won't teach you anything about how you *ought* to speak English. We'll leave that up to the language mavens and your own good judgment.

With those preliminaries completed, onward to our first topic: the sounds of English.

Appendix: Basic Grammatical Terms

Although this book is intended for people with no background in linguistics, I have assumed that most of you will know terms such as "noun," "verb," "subject," "suffix" and "prefix" already, or at least have a general idea of how they are used. Often these terms are used more generally or loosely in everyday speech than we will be using them here, so here are some rough-and-ready definitions and a few problem sets to help cement your familiarity with a few basic terms. These definitions also show up in the glossary at the end of the book, but you should be sure you understand them fully now, before reading further:

16

Affix   A covering term for both suffixes and prefixes.

Parts of speech   Parts of speech are also often called *syntactic categories* – just as we can say things like "The part of speech of *dog* is 'noun'.", we could say "The syntactic category of *dog* is 'noun'." All words have a part of speech – sometimes more than one. Here we'll look at just a few of the most basic; for more discussion, see Chapter 6.

Nouns are often defined as "people, places or things," and verbs as "actions, states or states of being," but this is definitely not adequate for our purposes. For instance, *attraction* is a noun, but it would be pretty crazy to call it a person, place or thing! Similarly, an *incantation* is an action, but it would be pretty crazy to call it a verb. Parts of speech are not defined by their meaning, but by their distribution – where they show up in a sentence, and what kinds of other words or affixes can go with them.

Noun   A listeme that:
- can be used as the subject of a sentence;
- can occur immediately following determiners (a.k.a. "articles") such as *the*, *one*, *some*, *any*, *this*, *a*, *many*, etc., or possessive pronouns such as *his*, *her*, *our* etc., with no other word in the phrase (see Chapter 6 for more discussion of these);
- can usually be marked with the plural suffix *-s*;
- can be modified with adjectives such as *pretty*, *happy*, *lucky*, *fortuitous*.

Verb   A listeme that
- can be marked for past tense (usually by putting *-ed* on it);
- can be suffixed with *-ing*;
- can be modified with words like *again*, *sometimes*, *often*;
- can occur immediately following auxiliaries, like *can*, *may*, *might*, *would*, *will*; also after negation (*not*, *can't*, *won't*), or the infinitive marker *to*.

Adjective   A listeme that
- can appear between a determiner and its noun, as in *the lucky cat*, modifying the noun;
- often ends in *-y*, *-ish*, *-ous*; often can be prefixed with *un-*;
- can be modified by words like *very* or *extremely*, as in *the very lucky cat*.

Adverb   A listeme that
- can modify a verb;

17

- often ends in *-ly*;
- can be modified by words like *very* or *extremely*, as in *extremely quickly*.

Prefix   A smaller-than-phonological-word-sized listeme that attaches to the beginning of another listeme: *un-* in *unhappy* is a prefix, *re-* in *refill* is a prefix, *dis-* in *disentangle* is a prefix.

Suffix   A smaller-than-phonological-word-sized listeme that attaches to the end of another listeme: *-s* in *dogs* is a suffix; *-ed* in *patted* is a suffix; *-ion* in *attraction* is a suffix.

Study Problems

1. Identify all the suffixes and prefixes in the following sentences. If you think something might be an affix but you're not sure, explain why you think it might be and also what it is that makes you unsure:
   a. It is often written that *antidisestablishmentarianism* is the longest word in the English language, but it isn't.
   b. Calamities are of two kinds: misfortune to ourselves, and good fortune to others. (Ambrose Bierce)
   c. If you want to know what God thinks of money, just look at the people he gave it to. (Dorothy Parker)
   d. It is difficult to produce a television documentary that is both incisive and probing when every twelve minutes one is interrupted by twelve dancing rabbits singing about toilet paper. (Rod Sterling.)
2. In each quote below, identify the nouns, verbs and adjectives that are employed. Again, if you think a particular word is being used as a noun, verb or adjective but are not sure, explain why you think it might be, and what it is that makes you unsure.
   a. The way to write American music is simple. All you have to do is be an American and then write any kind of music you wish. (Virgil Thompson)
   b. Do not, for one repulse, forego the purpose that you resolved to effect! (William Shakespeare, *The Tempest*)
   c. Men who are unhappy, like men who sleep badly, are always proud of the fact. (Bertrand Russell)
   d. It is a common delusion that you make things better by talking about them. (Dame Rose Macaulay)

3. Identify whether each of the following words is a noun, verb, adjective or adverb, Some belong, or can belong, to more than one part of speech. For each word, write a sentence in which you illustrate it being used as that part of speech. If you identify it as having two or more parts of speech, write two or more sentences, one illustrating each part of speech you think it has:

   publicly, love, government, bank, take, smart, sympathy, realistic, particularly, always, maturity, shelter, elegant, smooth, fast.

4. Two sets English pronouns, the *object* pronouns and the *possessive* pronouns, are given below:

	Object pronouns		*Possessive pronouns*	
1st	me	us	my	our
2nd	you	you	your	your
3rd	him/her/it	them	his/her/its	their

   Now consider the following subset (1st and 2nd person forms only) of another group of English pronouns, the *reflexive* pronouns:

	Reflexive pronouns	
1st	myself	ourselves
2nd	yourself	yourselves

   a. These reflexive pronouns are made up of two parts. What are they? Describe them using the name of the appropriate family of pronouns given above.
   b. State your description of the reflexive pronouns as a rule: "To create a reflexive pronoun, put a _____ pronoun together with the noun _____."
   c. Following your rule, what should the (four) English 3rd person reflexive pronouns be?
   d. What are the actual 3rd person reflexive pronouns in your dialect of English?
   e. Some dialects of English use the reflexive pronouns you created in (c), but they are not the forms used in Standard American English. What would a prescriptive approach to language have to say about the forms in (c)? What about a descriptive approach? Which dialect of English is more "logical" in its

formation of reflexive pronouns? Describe some prescriptive attitudes to English that you have encountered, discussing their pros and cons.

Further Reading

On modern linguistics and what it's about:

Pinker, Steven (1994) *The Language Instinct*. New York: W. Morrow and Co.

On what modern linguistics is *not* about:

Bauer, Laurie and Trudgill, Peter (1998) *Language Myths*. London: Penguin Books.

On Saussure and the arbitrariness of the sign:

Holdcroft, David (1991) *Saussure: Signs, System, and Arbitrariness*. Cambridge: Cambridge University Press.

Saussure, Ferdinand de ([1916] 1959) *Course in General Linguistics*. Charles Bally and Albert Reidlinger, eds. Wade Baskin, trans. New York: Philosophical Library.

General useful reference:

Crystal, David (1995) *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.

Huddleston, Rodney D. and Pullum, Geoffrey K. (2002) *The Cambridge Grammar of English*. Cambridge: Cambridge University Press.