

1

Introduction

What This Book Tries to Teach You

In the USA, and (I'm told) in some other English-speaking countries, grammar is no longer taught in schools; even classes in modern foreign languages sometimes try to teach with as little reference to formal grammar as possible. That is unfortunate for the following reason.

When you first start learning a new language, the number of new words you have to memorize in a hurry seems overwhelming; but actually the vocabulary is not the hardest part of a language to learn. True, most people need many dozen repetitions (in context) before they begin to recognize a foreign word automatically, but mere repetition will do the trick: if you put in the time and concentrate on what you're doing, you'll learn the words eventually.

However, it's much more difficult to learn the *structure* of a language – that is, its grammar – just by listening to it for hours. It wasn't always so. Up until the time you were about 7 years old, you really could learn a new language's grammar just by listening to it. Learning a language that way is called *native language acquisition*; any language acquired that way is a *native language*.

2 Introduction

You have at least one native language; many millions of people have two or three, and at least a few million have four or more (although few of those people live in the USA or in other English-speaking countries). If a language is native to you, you have a *perfect*, but *unconscious*, knowledge of its structure, and you can use it automatically without any conscious control.

Unfortunately, languages can be learned natively only in childhood; for reasons that we still don't understand, it becomes harder and harder to learn a new language that way late in childhood, and at puberty the "window" for native language acquisition closes for good. So what do you do if you need or want to learn a new language after that? Some educational systems simply keep plugging away with direct exposure, figuring that that's the best you can do. For a few students – gifted language learners, especially those that already know two or three languages – that works well, but for most students it takes a long time and still yields only modest results.

Fortunately you can improve on the "direct method" by *combining* it with a knowledge of how languages *in general* are structured; if you can get explicit instruction in the grammar of the language you're learning, the results are better still. In effect, since you no longer have a subconscious ability to learn language structure natively, you're substituting a conscious knowledge of language structure in general, or of the structure of your target language in particular.

In Europe they know that and act on it; that is one reason so many Europeans speak multiple languages (typically including English) fluently. Classicists, too, understand the necessity of explicitly teaching grammar. Professional linguists use the same techniques to learn new languages rapidly. Those three communities are proof that combining the teaching of formal grammar with extensive practice really works. The hours of direct exposure are still necessary, of course; advocates of the direct method are not wrong about that. The trick is to approach all that practice with as much advance information about the structure as possible, so that you can understand what you're hearing in the shortest possible time.

This book is intended to provide basic information about language structure so as to make your task in learning a foreign language easier. I haven't told you everything that's known about any topic – not even everything that I know. The aim is to get just the basics into your head. I have deliberately avoided introducing any specific modern linguistic theory; this book is "pretheoretical" (unlike, for example, Soames and Perlmutter 1979 or Gelderen 2002). That is not because I have any doubts about the advantages of modern linguistic theory;

after all, real progress in the study of syntax began with Chomsky, and we have learned an enormous amount from generative and post-generative work in linguistics. But a pretheoretical approach is simpler, and I have tried to keep this simple. I have also kept the discussion as brief as I could, since too much verbiage is bound to bore people. An unfortunate result of that strategy is that the book can't just be read cover to cover, because most people can't assimilate the material that fast. There are exercises at the end of many lessons to help you work into the material, but you (or your instructor) should also spend time trying to find further examples in a language (or languages) of interest to you; that will make the grammatical concepts easier to remember and use.

You should try to internalize the important points listed at the end of many of the lessons. Repetition helps much more than intensive cramming.

What This Book Expects from You

I assume that you really want to find out about natural human language structure so as to better understand the language you're learning and make it easier to learn. If that's true, you should be willing to learn a fair amount of unfamiliar but not-too-hard technical material, provided that I can explain it and present it to you at a reasonable pace. There really isn't any alternative; language structure is the way it is, and you have to take it on its own terms to understand it.

If you dread grammar the way some people dread math, or if your brain shuts down when it's presented with something in a foreign language, you need to force yourself to engage with the material; you'll soon find that it's more accessible and less intimidating than you had thought. Remember, millions of people have learned basic grammar.

If you're afraid of making mistakes, summon your courage and choke down the fear, because in language learning *making mistakes and having them corrected is the main way you learn*. Your instructor should understand that and should cut you lots of slack.

If you find grammar dry and boring, go back to direct exposure without explicit grammar – but it will take you longer to get where you want to go.

It's important to realize that the grammatical terms introduced (and there are a lot of them) are not jargon intended to exclude the uninitiated; they're tools for

4 Introduction

discussing the details of language structure. If we didn't use them, we'd have to describe a familiar concept every time we needed to refer to it, and technical discussion would become practically impossible. (The same is true in every technical discipline, above all in the "hard" sciences.) There is an extensive glossary at the end of the book; you should refer to it whenever you need to. But you should also try to remember the terms that keep appearing on a regular basis.

It's also important to realize that, with some basic knowledge of grammar, sharp powers of observation, and patience, it's possible to make sense of the structure of any language you encounter, no matter how unfamiliar it is. Moreover, looking at the grammars of several languages at once can actually be easier than tackling them one by one, because languages are enough alike that the structure of one will help illuminate the structure of another. Some of the exercises ask you to look at data from languages you have (probably) never seen before, which have not been discussed in the text, and answer questions about them. You will find that you can make sense of them, working with the English translations given and bearing in mind that you can usually cut a linguistic form up into its component parts.

If you can get interested in language structure, encountering new languages can be a great deal of intellectual fun.

Mastering the material in this book will require substantial effort, but in the long run it will pay off by making the task of learning a language easier.

The Structure of This Book

The universal concepts of language structure will be introduced in naturally connected groups. Each will be illustrated by examples from (up to) six languages: English, Spanish, Classical Latin, Biblical Hebrew,¹ Navajo, and Mandarin Chinese.² I will usually give the Mandarin examples both in pinyin and in (simplified) characters, for the benefit of readers who are used to characters; examples in

¹ I have occasionally used words from later stages of Hebrew when no word for a particular concept is attested in the Bible.

² There are several Chinese languages, all obviously related to but not mutually intelligible with one another. Mandarin is the majority language in China and the official language; it is usually what people mean when they say "Chinese."

languages other than English will sometimes be accompanied by notes or a word-for-word gloss in order to make their (relatively unfamiliar) structures clearer.

These languages were chosen because English, Spanish, and Mandarin are important world languages widely spoken in the USA and elsewhere; Latin and Biblical Hebrew are historically important in Western culture; and Navajo, which is among the Native American languages of the USA that are still viable, is superficially very different in structure from all the others examined here.

Together these languages give a fairly good picture of the range of possibilities in human language. Latin has a significant amount of INFLECTION; that is, its verbs, nouns, and adjectives appear in different forms depending on their relationships to other words in the sentence. Spanish is a direct descendant of Latin and therefore resembles it in many details, but it has simplified the inflectional system. English is distantly related to Latin and Spanish, so its grammatical categories are similar, but it has simplified the inherited inflectional system even more. The remaining languages are not demonstrably related to those three, nor to each other. Biblical Hebrew has an inflectional system roughly as complex as that of Latin but very different in detail; Navajo has a more complex inflectional system than either; Chinese languages have no inflection at all.

Of course, you might be studying some other language. Your instructor will be able to provide you with relevant examples from that language. You could read only the sections about English and your target language (if it is one of those employed here); after all, that's the material you'll need to remember. But if your main goal is to get a good idea of what language structure is like, then read about all the languages, even though you won't be able to remember everything.

In the first twelve lessons and in the fourteenth (on the English verb system), new terms will be in SMALL CAPITAL LETTERS. In subsequent lessons, not only new terms, but also those that have been introduced in lessons of less general interest (such as those on the verb systems of languages other than English) will be introduced in the same way. (In Lessons 27 and 28 most technical terms are not marked, since they occur only in those lessons.) Throughout the book, examples will be in *italics*; translations will be 'between single quotation marks'; and ungrammatical examples – sentences that you just can't say – will be preceded by an asterisk (*). When I give word-for-word translations into English (in order to make examples in other languages easier to understand), I will put the translations of inflectional markers in SMALL CAPS.

Practical Notes on Pronunciation (for the Interested and Curious)

The sounds of a language can be described adequately only by scientific phonetics and phonology, and we will take a brief look at that approach in Lesson 27. But in the meantime you might want at least a bit of information about how the example sentences in these lessons are pronounced. It's not strictly necessary, so you can skip this section if you like; you can always refer back to it if something puzzles you.

To begin with, you should be aware that English spells its vowels unlike any other language on the planet; the vowels themselves are not particularly strange, but the way they are spelled is bizarre. Therefore, if you pronounce written material in any other language the way you would in English, you are certain to be wrong. We need to start from a better reference point.

Spanish

Spanish provides a much better reference point. Like hundreds of other languages, Spanish has a simple system of five vowels: *i e a o u*. You can approximate their pronunciation in English by pronouncing *i* as “ee”, *u* as “oo”, and the others as “eh, ah, oh”. A better approach would be to listen to a native speaker pronouncing them. Audio files can be found at www.phonetics.ucla.edu/course/chapter1/vowels.html; you will see a chart with all the vowel symbols that phoneticians use. Click on each of the symbols that spell the normal Spanish vowels and you will access an audio file with a pronunciation of that vowel (not necessarily from Spanish, but in any case from a language with very similar vowels). The vowels of the other languages in our sample will be described below by reference to the Spanish vowels.

Consonants are somewhat less of a problem: many symbols have *approximately* (though not always exactly) the same pronunciation in numerous languages. Here are some notes on the most important points of Spanish pronunciation for a speaker of English to bear in mind.

Spanish has two r-sounds. At the beginning of a word *r* is heavily trilled; that is, the passage of air through the mouth makes the front part of the tongue hit the top of the mouth just behind the teeth several times in rapid succession. The same trill is spelled *rr* between vowels. In most other positions a single *r* is just a light tap of the tongue-tip against the top of the mouth just behind the teeth. The two Spanish *r*'s need to be distinguished, because the difference between them can make a difference in meaning; *pero* 'but' is a different word from *perro* 'dog'. The Spanish consonant *ñ* is approximately like the *ny* in English *canyon*; in Spain *ll* is approximately like the *lli* in English *million*, but in the Americas *ll* is pronounced exactly like *y*. The consonant *s* is always voiceless,³ as in English *hiss*; there is no consonant like English *z* in Spanish. In the Americas the Spanish consonant spelled *c* before *e* and *i*, but *z* in other positions, is pronounced exactly like *s*; in Spain it is pronounced like the *th* of English *thin*. The spelling of k-, g-, and h-sounds is more complicated:

approximate Engl. equivalent	before <i>e</i> and <i>i</i>	elsewhere
<i>k</i>	<i>qu</i>	<i>c</i>
<i>g</i> (as in <i>get</i>)	<i>gu</i>	<i>g</i>
<i>h</i>	<i>g</i>	<i>j</i>

The written letter *h* is not pronounced at all in Spanish; for instance, *hombre* 'man' is pronounced as if it were spelled "ombre".

Latin

The Latin vowel system is like that of Spanish, but with an additional complication: each of the five vowels occurs both long and short. Long vowels are written in textbooks with a mark ¯ over them. The long vowels take about twice as much time to say as the short ones, so that from an English point of view the long vowels sound drawled and the short vowels can sound clipped. For instance, the first vowel of *mālus* 'mast' is much longer than the first vowel of *malus* 'bad', although otherwise they sound very similar – and as this example shows, a difference in

³ See Lesson 27 for a description of voicelessness.

8 Introduction

vowel length often makes a difference in meaning. Latin also has diphthongs, in which you begin with one vowel and slide into another; the two common ones are *ae*, pronounced roughly like the diphthong in English *height* or *chai*, and *au*, pronounced like the diphthong in English *house*.

For Latin consonants only a few notes are necessary. As in Spanish, *s* is always voiceless. Latin *c* is always pronounced like English *k*; *g* is always “hard,” as in *get*; *v* is pronounced like English *w*; *qu* is “k” + “w”, as in English.

Navajo

The Navajo vowel system is similar to the Latin system, with two major differences: there are only four basic vowels (there is no *u*), and in addition to being either long or short, vowels can be nasalized or not. Nasalized vowels are like the French vowels written *an*, *in*, *on*; in Navajo they are written with a hook beneath them, and long vowels are written double. For instance, in *háádéé'* ‘from where?’ both vowels are long, and the second vowel is nasalized. Navajo also has a system of tones, according to which the pitch of the speaker’s voice rises and falls; each vowel is spoken on a (relatively) high or low pitch. High tone is marked with an accent; thus both syllables in the word just cited have high tone (i.e. high pitch relative to the syllables around them in an utterance).

Many Navajo consonants and combinations of consonant-letters express approximately (though often not exactly) the same sounds as in English. Navajo *zh* spells the middle consonant of English *vision* or *measure*. Navajo *l* is a voiceless l-sound, rather like a whispered *l*, but noisier; such a sound does not occur in English, but it does in Welsh, where it is spelled *ll*. (To find it on the UCLA phonetics website, look for the phonetic symbol [ɭ].) There is also a consonant spelled *gh*, which is like English (hard) *g* except that the tongue never actually reaches the roof of the mouth; it is exactly like Spanish *g* between vowels, but whereas in Spanish that is simply a special pronunciation of *g*, in Navajo it is a separate consonant. The *glottal stop*, spelled with an apostrophe (’), is a full consonant in Navajo; it is made by closing off the throat for a split second and then releasing it, as in the middle of the American English interjection *uh-uh* (meaning ‘no’). Finally, Navajo has *glottalized* or *ejective* consonants, spelled *t’*, *tl’*, *ts’*, *ch’*, *k’*; as you might expect, they are made by producing the consonant more or less simultaneously with a glottal stop.

Mandarin

Mandarin, like Spanish, makes no distinction between long and short vowels. It uses the same symbols as Spanish, plus *ü*, which is a high front round vowel like French *u* or German *ü*. But the pronunciation of Mandarin vowels shifts dramatically depending on what other sounds they are adjacent to. For instance, *e* represents the sound spelled “eh” in English when it immediately follows *i* or *y*, but in most other positions it sounds like English “uh”; *u* is usually pronounced as in Spanish (= English “oo”), but after certain consonants it is instead pronounced like *ü*; and so on. The details are too complex to go into in these notes. More importantly, Mandarin is a tone language, like Navajo; but Mandarin has four tones, distinguished by marks over the vowel when spelled in the Roman alphabet. Using the vowel *a* as an example, *ā* has a high tone; *á* has a rising tone, usually somewhat drawled; *ǎ* has a low tone that dips in the middle; and *à* has a sharply falling tone and usually sounds clipped. The tone of a syllable is an integral part of it: if you change the tone, you produce a different syllable with a different meaning. For instance, *chū* means ‘to come/go out’, *chú* ‘to divide’, *chǔ* ‘to get along with (someone)’, and *chù* ‘to touch’.

The way some Mandarin consonants are spelled is very different from the English system. Mandarin *s*, *sh*, and *ch* are much like English; but *c* is “ts”, *z* is (roughly) “dz”, and *zh* is approximately like English “j”. In addition, there are three consonants that have no English equivalents: *x* is like “sh” pronounced simultaneously with a “y” sound; in the same way, *q* is like “ch” pronounced simultaneously with “y”, and (Mandarin) *j* is like (Mandarin) *zh* pronounced simultaneously with “y”. Distinguishing *x*, *q*, *j* from *sh*, *ch*, *zh* respectively takes a good deal of practice; you should get a native speaker to help you if you can.

Biblical Hebrew

What the Biblical Hebrew vowels were like depends on what era you’re talking about. The traditional system of transcription that I use attempts to recover the pronunciation of about the fifth century BCE; it is possible that the distinction

10 Introduction

between long and short vowels had been lost by then, but if so, the transcription simply reflects a somewhat earlier stage of the language. Most of the system resembles that of Latin, but in addition there is a vowel *a*, called “schwa,” which is the indistinct vowel heard in the first syllable of English *about*, the last syllable of *sofa*, or the middle syllable of *terrible* or the last two of *tenement*. (Any vowel symbol can be used to spell schwa in English!) The vowels written small above the line (e, a, o) are hypershort versions of the corresponding full vowels.

Hebrew *š* is very much like English “sh”; scholars now believe that *ś* (*śīn*) was actually pronounced like Navajo *l* (see above), although it eventually became identical with *s* (*sāmek*). In fact there are many different ways to pronounce the system of Hebrew consonants; for instance, you can pronounce Biblical Hebrew like Modern Hebrew, and many people do (just as Italians pronounce Latin – the ancestor of their language – like Italian). As in other such cases, the ancient pronunciation was somewhat different; the following notes describe what scholars think it was like in the fifth century BCE. The consonants *b*, *d*, *g* (as in *get*), *k*, *p*, *t* were pronounced roughly as in English in many positions, but after vowels they were lenited (“softened”). The lenited pronunciations and the transcriptions used in this book are:

basic consonant	lenited consonant	rough English equivalent
<i>b</i>	<i>b̄</i>	<i>v</i>
<i>d</i>	<i>d̄</i>	<i>th</i> as in <i>either</i>
<i>g</i>	<i>g</i>	(no equivalent; = Navajo <i>gh</i>)
<i>k</i>	<i>k̄</i>	(no equivalent; = <i>ch</i> in German <i>Bach</i>)
<i>p</i>	<i>p̄</i>	<i>f</i>
<i>t</i>	<i>t̄</i>	<i>th</i> as in <i>ether</i>

Like Navajo, Hebrew had a glottal stop, here transcribed ’; and scholars believe that the consonants transcribed *ṭ*, *ṣ*, *q* (*ṭēṭ*, *ṣādē*, *qōp̄*) were glottalized, corresponding to Navajo *t’*, *ts’*, *k’*.⁴ Finally, Hebrew had two pharyngeal consonants,

⁴ It is not certain that *ṣ* was a glottalized affricate [tʰ] rather than a glottalized fricative [sʰ] in the fifth century BCE, but it was certainly an affricate in the Middle Ages and is one in Modern Hebrew. I am grateful to Aaron Rubin for helpful discussion of this point.

pronounced by pulling the tongue back to narrow the air passage behind it: the voiceless (roughly, whispered) one is transcribed *ħ* (*ħēṭ*), the voiced one with a reversed apostrophe ‘ (*‘ayin*). They are extraordinarily difficult for a speaker of English to pronounce, but they are common in Arabic, which is fairly closely related to Hebrew. If you can’t find a speaker of Arabic to help you, examples can be found by surfing the UCLA phonetics website (referenced above) under the symbols [ħ] and [ʕ].

One final peculiarity of Hebrew spelling should be noted here. Even in the fifth century BCE, not all the consonants that were written were pronounced; in particular *h* at the end of a word or before another consonant was usually silent, and the glottal stop ‘ was also very often silent in those positions. There were also some combinations in which *y* and *w* were silent. I have *not* transcribed the silent letters of Biblical Hebrew, with the result that my transcription of the language diverges somewhat from the usual spellings.