# Introduction to Linguistic Science

Mark Hale
Concordia University
Montréal, Québec

Madelyn Kissock
Oakland University
Rochester Hills, MI

2001

# Contents


**1 Introduction** **5**
1.1 What is 'linguistics'? . . . . . 5
1.2 'Language' vs. languages . . . . . 6
1.3 Where is 'language'? . . . . . 9
1.4 What is a linguistic system? . . . . . 12
1.5 What linguistics is *not* about . . . . . 14
1.6 Exercises . . . . . 15

**2 Phonetics** **17**
2.1 Introduction . . . . . 17
2.2 Some further implications of Figure 2.1 . . . . . 20
2.3 Articulatory Phonetics . . . . . 22
2.3.1 Manners and Places of Consonantal Articulation . . . . . 24
2.3.2 Some Further Consonantal Details . . . . . 35
2.3.3 Vowels . . . . . 35
2.4 Summary of relevant IPA symbols . . . . . 38
2.5 Exercises . . . . . 41

**3 Phonology** **46**
3.1 Further Phonetic Details . . . . . 46
3.2 The Phoneme . . . . . 48
3.3 Minimal Pairs & Allophones . . . . . 50
3.4 More English allophony . . . . . 53
3.5 Rule ordering . . . . . 55
3.6 Some Allophony in other languages . . . . . 58
3.7 Universal Properties and Language-Specific Properties . . . . . 59
3.8 Summary . . . . . 60
3.9 Hints on solving phonology problems . . . . . 61

3.10 Exercises . . . . . 64

**4 Morphology** **72**
4.1 Morphemes . . . . . 72
4.2 Morphemes and Allomorphs . . . . . 82
4.3 Productivity . . . . . 88
4.4 Hierarchical Structure in Morphology . . . . . 89
4.5 The ‘cranberry’ problem . . . . . 92
4.6 Summary . . . . . 93
4.7 Exercises . . . . . 94

**5 Syntax** **101**
5.1 Introductory remarks . . . . . 101
5.2 ‘Grammaticality’ . . . . . 102
5.3 The simple sentence in English . . . . . 104
5.4 Categories . . . . . 107
5.5 Constituency . . . . . 112
5.6 Structural Ambiguity . . . . . 116
5.7 Some Problems . . . . . 118
5.8 D-structure, S-structure, and Transformations . . . . . 121
5.9 Yes-No Questions . . . . . 124
5.10 Summary . . . . . 131
5.11 Exercises . . . . . 131

**6 Semantics** **134**
6.1 Introduction . . . . . 134
6.2 Lexical Semantics . . . . . 135
6.3 Interpretation of Sentences . . . . . 139
6.4 Summary . . . . . 144
6.5 Exercises . . . . . 144

**7 Sociolinguistics** **146**
7.1 Introduction . . . . . 146
7.2 Traditional Dialectology . . . . . 148
7.3 Problems with Traditional Dialectology . . . . . 151
7.4 The Sociolinguistic Revolution . . . . . 152
7.5 Standard and Non-Standard . . . . . 156
7.6 Gender and Linguistic Variation . . . . . 158

7.7 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 160
7.8 Exercises . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 162

**8 Historical Linguistics 164**
8.1 Language Change . . . . . . . . . . . . . . . . . . . . . . . . . 164
8.1.1 Phonological Change . . . . . . . . . . . . . . . . . . . . 167
8.1.2 Morphological Change . . . . . . . . . . . . . . . . . . . 169
8.2 Comparative Reconstruction . . . . . . . . . . . . . . . . . . . 171
8.3 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 189
8.4 Exercises . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 189

**9 Language Acquisition 193**
9.1 The Earliest Stages . . . . . . . . . . . . . . . . . . . . . . . . 194
9.2 Later Stages . . . . . . . . . . . . . . . . . . . . . . . . . . . . 197
9.3 Theories of Acquisition . . . . . . . . . . . . . . . . . . . . . . 199
9.4 The Critical Period Hypothesis . . . . . . . . . . . . . . . . . . 202
9.5 Summing Up . . . . . . . . . . . . . . . . . . . . . . . . . . . . 203

# List of Figures


2.1 The mind vs. the world. . . . . . . . . . . . . . . . . . . . . . 18
2.2 The components of articulation. . . . . . . . . . . . . . . . . . 24
2.3 Places of articulation. . . . . . . . . . . . . . . . . . . . . . . 26
2.4 Stops. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 30
2.5 Fricatives and affricates. . . . . . . . . . . . . . . . . . . . . 33
2.6 The vowel space. . . . . . . . . . . . . . . . . . . . . . . . . 36
2.9 The primary consonants of English. . . . . . . . . . . . . . . . 38
2.11 Examples for ‘unusual’ IPA symbols. . . . . . . . . . . . . . . 39
2.7 The simple vowels of English. . . . . . . . . . . . . . . . . . . 40
2.8 The diphthongs of English. . . . . . . . . . . . . . . . . . . . . 40
2.10 The simple vowels of English. . . . . . . . . . . . . . . . . . . 40

7.1 Percentage of speakers with two /r/’s in casual ‘fourth floor.’ . 154
7.2 Percentage [ɪŋ] forms produced. . . . . . . . . . . . . . . . . . 159
7.3 Number [ɪŋ] and [ɪn] Forms Produced By Boys and Girls Aged 3 to 10. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 160

# Chapter 1

# Introduction

## Contents


1.1 What is 'linguistics'? . . . . . . . . . . . . . . . . . 5
1.2 'Language' vs. languages . . . . . . . . . . . . . . . 6
1.3 Where is 'language'? . . . . . . . . . . . . . . . . . . 9
1.4 What is a linguistic system? . . . . . . . . . . . . . 12
1.5 What linguistics is *not* about . . . . . . . . . . . 14
1.6 Exercises . . . . . . . . . . . . . . . . . . . . . . . . . 15


## 1.1 What is 'linguistics'?

Linguistics is the study of one aspect of human knowledge, specifically, knowledge of language. Because linguistic science concerns human knowledge, it falls under the scope of the broader field of Cognitive Science. Other disciplines which fall under Cognitive Science include those that study auditory perception, visual perception and motor coordination. As is usual in scientific enterprises, 'language' has a very specific sense in linguistics – one which does not match that found in everyday usage. Since we will be using 'language' in its technical, linguistic sense, it is important at the outset to be clear about what precisely the term means. In this context – an introduction to a scientific pursuit – this precision has the added benefit of giving you some idea what the subject matter of the discipline is. This chapter is intended to be helpful in this regard.

Physics is the study of the physical universe. Nevertheless, we do not normally see physicists wandering around measuring dogs, beards, or mailboxes – although all these things are part of the physical universe. Physicists attempt to develop general rules for the structure and processes of the physical universe from which the properties of specific entities – a dog, a beard, or a mailbox – follow. Similarly, the object of study in linguistics is not your language or Fred's language or any specific instantiation of 'language', but rather the general principles which govern all languages. Specific individual languages, like specific entities in the physical universe for a physicist, are intended to follow from this general model of language.

Contrary to widespread everyday usage, then, a linguist is NOT someone who knows a lot of languages.[1] It is true, however, that a primary source of data for the general theory of the nature of linguistic knowledge must necessarily be actually observed linguistic systems. Therefore, linguists generally know a certain amount about the structures and processes found in a variety of languages; however, the knowledge that linguists have about languages (other than their native language) is what we might call **explicit** or **conscious** knowledge. By contrast, the knowledge that a native speaker has is **tacit** or **unconscious** knowledge. (If it were not, the job of the linguist would be quite trivial.) The linguist's explicit knowledge about the structures and processes found in the linguistic systems they study does *not* give them the ability to speak these languages. If you have ever studied a second language in some relatively formal setting, you know how inadequate conscious knowledge of the grammar of a language is as a basis for speaking. It is only when your knowledge becomes *unconscious* that you actually feel you can speak a language.

## 1.2 'Language' vs. languages

One of the most frequent uses of the term 'language' in everyday speech is in such collocations as 'the English language,' 'the French language,' 'the Swahili language,' etc. These notions of 'language,' which we will call 'sociopolitical,' turn out to have no clear referent and thus provide no meaningful basis for scientific inquiry.

[1]Saying 'John is a great linguist, he knows 10 languages' sounds to a linguist like saying 'Mary is a great physicist, she's good at carpentry.' (Carpentry involves manipulation of entities in the physical universe.)

We will give two examples of the difficulties which confront one attempting to develop a coherent notion of 'the English language.' The first is literary (in spite of the fact that written language does not fall under the purview of linguistic science in the view of most linguists). The Modern Library Board recently released a list of the top 100 of "what it considers to be the best novels published in the English language since 1900." James Joyce's *Finnegans Wake* is listed as number 77, meaning, among other things, that this board of specialists believe it is written in 'the English language' (and that they liked it). A typical passage from *Finnegans Wake* is this nice little scene, early in the novel (pp. 15-16):

> In the name of Anem this carl on the kopje in pelted thongs a parth a lone who the joebiggar be he? Forshapen his pigmaid hoagshead, shroonk his plodsfoot. He hath locktoes, this shortshins, and, Obeold that's pectoral, his mammamuscles most mousterious. It is slaking nuncheon out of some thing's brain pan. Me seemeth a dragon man. He is almonthst on the kiep fief by here, is Comestipple Sacksoun, be it junipery or febrewery, marracks or alebrill or the ramping riots of poutiose and froriose. What a quhare soort of a mahan. It is evident the michindaddy. Lets we overstep his fire defences and these kraals of slitsucked marrogbones. (Cave!) He can prapsposterus the pillory way to Hirculos pillar. Come on, fool porterfull, hosiered women blown monk sewer? Scuse us, chorley guy! You tollerday donsk? N. You tolkatiff scowegian? Nn. You spigotty anglease? Nnn. You phonio saxo? Nnnn. Clear all so. 'Tis a Jute. Let us swop hats and excheck a few strong verbs weak oach eather yapyazzard abast the blooty creeks.

If a linguist attempted an analysis of our instantiation of knowledge of language, which most people would call 'English,' and had also to account for the above data as 'the English language,' he or she would certainly fail to develop a coherent account of either our knowledge or the knowledge of Joyce.

To take a more mundane example, here is a little story from Yorkshire (taken from Arnold Kellett, *Basic Broad Yorkshire*, Smith Settle, 1992). We have added a translation.

> 'Thoo's getten poison i' thi' sistren, that's why thoo's belly-wark,' says t' docther. 'Thoo mun a thi teeath oot.'
>
> 'What?' Ah says, 'All on 'em?'
>
> 'Aye,' he says. 'ivvery yan'... So Ah took 'em oot, an laad em on t' table.

> 'You're getting poison in your well, that's why you have a belly-ache,' says the doctor. 'You have to have your teeth extracted.'
>
> 'What?' I say, 'All of them?'
>
> 'Yes,' he says, 'every one.'... So I took them out and laid them on the table.

While the attempt to represent this way of speaking using a modified version of the traditional English writing system leaves much unclear (we will see how linguists deal with this problem soon), what is clear is that *our* language and this one are quite different. Again, any attempt to develop a scientific analysis of our speech which must also deal with this evidence will fail to provide insightful analysis of either linguistic system.[2]

What these examples show is that the notion 'the English language' does not provide the linguist with a sufficiently coherent body of data to subject to meaningful scientific analysis. To understand the differences between our language and that of Joyce, that of Yorkshire, that of an Alabama farmer, or that of a life-long resident of Melbourne, Australia we cannot lump these quite diverse linguistic systems into one big 'English language' bucket. Indeed, it turns out that upon sufficiently close examination, the language which each of us uses is different from the language of virtually everyone else. Even the notion 'the language which I use' turns out to be rather underdifferentiated, for scientific purposes.

A cautionary note must be added at this point. Linguists continue to use such terms and phrases as 'English', 'in Tzotzil', and 'some languages have... ' While this is both misleading and inaccurate, it is convenient, and it is done *with the understanding* that these entities do not actually exist in the relevant linguistic sense.

[2]By the way, once we recognize that the Yorkshire linguistic system and ours are different in important ways, if we want to keep calling one of them 'English,' it is probably we who are in trouble: Yorkshire is in England, after all...

## 1.3 Where is 'language'?

The foregoing discussion reveals a serious shortcoming in traditional notions of what 'language' is. 'Language' is generally treated as something 'out there,' in the world. The meaning of words is thought to be found by consulting a dictionary, the 'correct' rules of grammar by looking in a grammar book. There is some notion that correct 'English' exists and that we are all attempting to attain this difficult target, some of us with more success than others.

Linguists, by contrast, recognize the important role that *all* linguistic systems play in the scientific pursuit of attempting to understand in general what a possible linguistic system is. Through an examination of the diversity of human linguistic systems, linguists have learned that there are no 'incorrect' linguistic systems, no 'primitive' languages, and, most importantly, that language, in the relevant scientific sense, is not something 'out there' but rather something 'inside' a human being – a piece of mental machinery. Each individual has a particular knowledge state with respect to language. As a result, there are at least as many 'languages' as there are humans.[3] A reasonable question at this point might be whether this reduces the possibility of scientific inquiry into knowledge of language to zero. If everyone's knowledge state were completely different, it would mean that no interesting generalizations about language (and consequently about human knowledge with respect to it) could be made. What linguists have observed, though, is that no matter how superficially different the languages are (in the sounds or words they use, for example), there are striking similarities among them whether those languages are types of Chinese, Telugu, Zulu or Pohnpeian.

Our goal, then, is to characterize what humans *actually* know with respect to language, not what society/books/educational systems state that they are *supposed* to know.[4] An important step in creating a model of human knowledge of language is to develop **descriptive grammars**. These contrast sharply with the **prescriptive grammars** that most people are familiar with from school. An example might make this clearer. If you were to observe our

[3]The inclusion of 'at least' is meant to account for individuals who are multilingual and/or multidialectal.

[4]We are not making a judgement about the validity of the latter. We are simply saying that what people are supposed to know, for whatever reason, is irrelevant to linguistic inquiry.

speech for some period of time, you would find that we say things like the following:

- I have to teach this morning and so don’t you.
- I didn’t do nothing all day yesterday.
- Since you never come to class, that’s all the higher a grade I can give you.

Instead of (in keeping with ‘grammar’ books):

- I have to teach this morning and so do you.
- I didn’t do anything all day yesterday.
- Since you never come to class, that’s as high a grade as I can give you.

There are several possible theories one could develop as to why we say one set of sentences, rather than the other. These include (among others):

1. We’re of well below average intelligence.
2. We’re lazy.
3. We’re of well below average intelligence and we’re lazy.
4. We’re undereducated.
5. We’re undereducated and of well below average intelligence.
6. We’re undereducated and we’re lazy.
7. All of the above.
8. We have different linguistic systems than that of the people who wrote the grammar books.

We are not really in a sufficiently impartial position to evaluate our degree of intelligence, sloth, and education (although we both hold Ph.D.'s), but we think there are pretty compelling reasons to believe that (8) provides the most productive hypothesis. Consider, for example, the following. In English grammar books, multiple negation ('I didn't do nothing all day yesterday.') is considered non-standard (i.e., incorrect). By contrast, in many linguistic systems, including those generally referred to as 'Slavic,' multiple negation is standard. Failure to use both negatives in such sentences would be taken as a clear sign of lack of intelligence, laziness, or lack of education. But how can it be laziness for one of us to use multiple negation *and* for speaker of a 'Slavic'-type language to fail to do so? Either one or the other must be the lazier option, and lazy people everywhere should opt for the same form. Instead, what counts as lazy, stupid, or lacking in intelligence is always related to some societal norm, the workings of which are quite independent of any linguistic system.

If we have different linguistic systems than those who wrote the grammar books, where do we have them? How did we get them? What are they? These are basic questions posed by linguistic science. The generally accepted answer to the first of these questions is not surprising: an individual's linguistic system is located in their mind. It is that component of the cognitive system which enables a person to parse (analyze) incoming speech and to produce linguistic output. The property of having human-type linguistic systems appears to be restricted to human-type minds. Although efforts have been made to teach various symbolic manipulation systems to higher apes, the fact remains that if we take any normal human infant and expose them to human speech, they will acquire a human-type linguistic system. This is not true, as far as can be determined, of apes, dogs, rats, mosquitos, or potatoes. It is uniquely true of humans, to our present knowledge.

The answer to the question of how we each ended up with the linguistic systems that we have is contained in many ways in the preceding paragraph: when we were infants, people spoke to us and around us. As in humans generally, this led to the development of a linguistic system in each of our minds/brains (commonly known as our 'native languages'). The process of how this happens will be the subject of considerable further discussion.

The process of language acquisition also explains to a large extent why most of our linguistic systems differ from those of the people who write grammar books. They were *not* the ones speaking to us when we were children. The people who were speaking to us when we were children said things like 'I didn't do nothing all day yesterday' and 'I have to go to school now and so don't you', so we say things like that.

## 1.4 What is a linguistic system?

This question will be the primary concern of the first two-thirds of this book. Some general considerations are addressed below, however.

The linguistic system is stored in the human mind – it is some type of knowledge. As pointed out earlier, each person's linguistic system is somewhat different from another's but, at the same time, there are striking similarities in human linguistic systems. In fact, the set of possible human linguistic systems appears to be highly restricted. These restrictions on human linguistic systems can be accounted for if we assume that innate (i.e., genetically given) knowledge restricts the types of linguistic systems humans can create. The genetically-determined portion of the system must necessarily be supplemented by data acquired from the environment. One way to answer the question of what 'language' is, in the linguistic sense, is to explore what one must know to know a language.

The traditional subfields of linguistics provide broad insight into this issue. The basic subfields are:

1. Phonetics & Phonology
2. Morphology
3. Syntax
4. Semantics

**Phonetics & phonology** concern themselves with the knowledge which underlies the ability to produce and analyze articulated output of the grammar. For example, if you were asked to make up a name for a newly created beard-trimming mechanism, you might call it a

- beard-wrangler
- zorp

- flirp

but you would almost certainly *not* call it a

- dnkli
- ngloopi
- [lyn] (=the usual pronunciation of French 'lune')

You appear to have some knowledge of what is a 'possible word' in your linguistic system. The first set of examples satisfies this knowledge, whereas the latter are in conflict with it.

Morphological knowledge is reflected in your ability to analyze words generated by linguistic systems into meaningful parts. For example, you know that 'cats' consists of two parts: a part that means 'cat' (that's the 'cat' part) and a part that means 'more than one' (that's the '-s'). You know that it is *not* the case that it consists of, e.g., three parts (a part that means 'domesticated animal,' a part that means 'feline' and a part that means 'more than one'). The internal structure of words is analyzed by the **morphology** of the linguistic system, which also produces internally-complex words.

Syntactic knowledge is knowledge about the structure of sentences. You know that

- you will not pass this class if you don't pay attention

is a possible sentence given your linguistic system. On the other hand,

- *not class this pass don't pay if will attention you you

is not.[5] Your linguistic knowledge, in particular the **syntax** of your linguistic system, allows you to see that this is just a list of words, not a sentence.

Semantics is the study of meaning. As such, its concerns are much broader than those of linguistics alone. However, some aspects of meaning may be

[5]Linguists mark strings of words which do not parse as sentences – so-called 'ungrammatical' strings – with an asterisk, as here.

linguistic in nature (the matter is quite difficult). It is certainly the case that the grammar plays a key role in one's ability to assign meaning to sentences that one hears and to produce sentences that express, at least in some vague way, the meaning one intends in some particular situation.

These four components make up what linguists usually call 'the grammar.' By this they of course do not mean the kind of prescriptive grammar found in grammar books, but rather a component of the cognitive system of a human – a mental organ, as it is sometimes called, which has the responsibility for producing and parsing linguistic information. We will study each of these 'modules' of the grammar.

## 1.5 What linguistics is *not* about

Before proceeding, we would like to point out certain topics that we are *not* going to talk about and why we are not going to talk about them. The first of these is communication. It is true that the human linguistic system is often (but not always) used for communicative purposes – so, for that matter, are Morse Code, semaphores, and picture-drawing, among other things. Note that the same message can be communicated using a variety of different systems, such as those mentioned above. Given this fact, the message is independent of the system (although it is true that some systems convey certain messages better than others). Linguistic inquiry of the type we have described here is an exploration of the properties of the linguistic system, itself, not the message or messages that are conveyed by it.

The second topic that we will not talk about is writing systems. Human knowledge of language existed long before writing systems developed and continues, for the vast majority of languages, to exist independently of such systems. All humans have a fully-developed linguistic system before they learn a writing system (if, indeed, they ever learn one).

## 1.6 Exercises

Circle the correct answer.

1. A speaker's knowledge of his/her language is:
   (a) explicit knowledge.
   (b) conscious knowledge.
   (c) tacit knowledge.
   (d) all of the above.
2. 'I'm not doing nothing today' is:
   (a) a string of words which is ungrammatical for all humans.
   (b) a string of words which is a grammatical sentence for some humans.
   (c) an illogical statement.
   (d) a fundamental truth.
3. 'Syntax' is the module of the grammar responsible for:
   (a) the structure of sentences.
   (b) the structure of complex words.
   (c) the sounds used in human languages.
   (d) that extra money you have to pay when you buy cigarettes.
4. Linguists examine various linguistic systems and develop
   (a) prescriptive grammars.
   (b) descriptive grammars.
   (c) dictionaries.
   (d) theories about which systems are best.
5. A 'linguist' is someone who:
   (a) can speak more than five languages.
   (b) can speak more than ten languages.
   (c) investigates the nature of human language.
   (d) speaks with proper grammar.

6. Select the words below which belong in this sentence: ‘The English language is ________ linguistic research.’
   (a) a critical component of
   (b) the major focus of
   (c) irrelevant to
   (d) one aspect of
7. ‘Phonology’ is:
   (a) the study of telephonic communication.
   (b) the study of telepathic communication.
   (c) the module of the grammar which allows one to produce and parse speech sounds.
   (d) You phonio saxo? Nnnn.

# Chapter 2

# Phonetics

## Contents


**2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . 17**

**2.2 Some further implications of Figure 2.1 . . . . . 20**

**2.3 Articulatory Phonetics . . . . . . . . . . . . . . . . 22**

2.3.1 Manners and Places of Consonantal Articulation . 24

2.3.2 Some Further Consonantal Details . . . . . . . . . 35

2.3.3 Vowels . . . . . . . . . . . . . . . . . . . . . . . . . 35

**2.4 Summary of relevant IPA symbols . . . . . . . . 38**

**2.5 Exercises . . . . . . . . . . . . . . . . . . . . . . . . 41**


## 2.1 Introduction

Given that what we linguists are interested in is a property of the human mind, we have a problem. The mind itself is clearly not directly observable.[1]

[1]There are new and exciting techniques for studying electro-chemical activity in the *brain*, such as PET scans and MRIs, however, we know virtually nothing about the relationship between the cognitive systems posited by scientists for the mind and their physical instantiation in brains. While recent developments may have moved this question from the domain of mystical speculation into that of potential empirical investigation, we need to have well-grounded theories in three domains: (1) the cognitive systems themselves, (2) the physical structures and processes of the brain, and (3) the relationship between (1) and (2). Linguistics is deeply involved in (1). At present (2) is rather poorly understood. (3) remains mysterious.

Linguists have a variety of tools which make it possible to transcend the difficulties created by the problem (which confronts all scientific inquiry to some degree). The figure below may help clarify why some of these tools are useful to linguists.

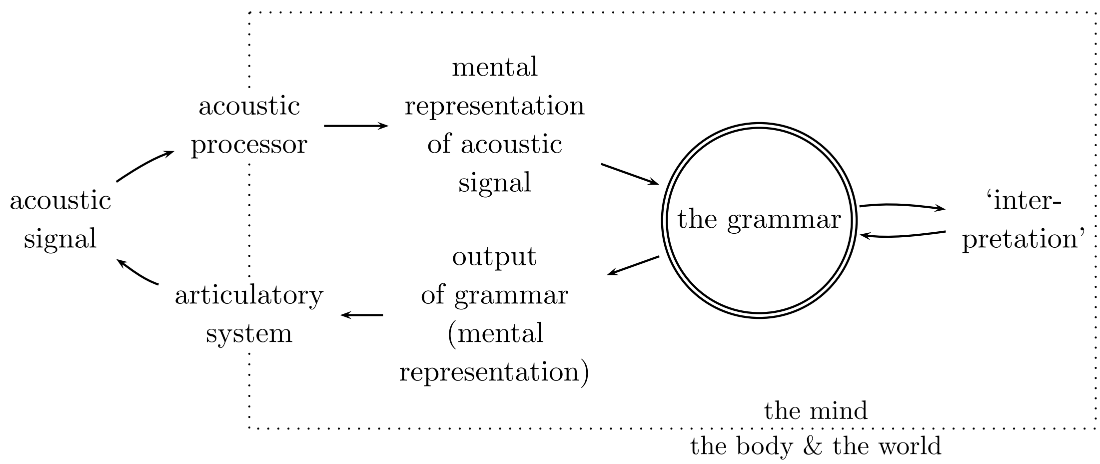



Figure 2.1: The mind vs. the world.

In this figure, we have attempted to show schematically the two functions of the grammar: that of generation of output and that of parsing (analysis) of input. The arrows leading *into* the grammar indicate the processing of input data, the arrows leading *away from* the grammar indicate the process of generation of spoken output. The dotted box is intended to reveal which aspects of these processes are *mental* as opposed to *physical*.

Examining the input path first, we can see that we start with an acoustic signal.[2] If this signal is within earshot of our schematic listener, it will enter his or her ear, be transformed through fairly complex mechanisms into neural impulses, and thus 'be heard' in some meaningful sense. Roughly, hearing a sound involves converting a physical set of 'sound waves' (variations in air pressure) into a mental representation. If the signal is human speech,

[2]Natural (human) language may use an acoustic or a visual signal as the medium. The structure of the argument is not affected by the medium. The technical details in the chapter are based on an acoustic signal rather than a visual/signed one, however. The basic principles upon which the linguistic system is built are the same for both acoustic/spoken and signed natural language.

this mental representation will be passed on to the grammar for linguistic processing. Assuming the signal is parsable by the grammar, the grammar will generate an 'interpretation' for the signal. Naturally, this interpretation is a mental representation, as well.

The line between physical and cognitive processes is found at that point at which the physical effects of receipt of a sound wave by the part of the physiology responsible for auditory processing are converted into a form which the rest of the cognitive system can operate on. The actual sound waves do not ricochet off the inside of the skull – they must be cognitively represented in order to be processed by higher cognitive systems. This is indicated in the figure by the fact that the mind/body line cuts through the 'acoustic processor' module.

If we now consider the *generation* of speech in this figure, we see that we start with some kind of semantic representation (an 'interpretation') which the speaker wishes to produce as speech. This semantic representation serves as the input to the grammar. The grammar acts as a processor which converts the semantic/input representation into an output representation. Note that the output representation is *not* part of the grammar – it is the product of the grammar's internal processing. This output representation is then passed on to the articulatory system of the speaker, which converts it into a set of muscle commands, implemented by the upper respiratory system. The result of this processing is a set of articulatory acts performed in response to the muscle commands. These acts will generate speech.

The line between physical and cognitive processes in speech generation is found at that point at which the mental representations generated by the grammar are converted into a set of physical signals sent to the articulatory organs. In Figure 2.1, this is indicated by the fact that the dotted line which separates the mind from the physical world cuts through the 'articulatory system' module.

Having gone through these issues in some detail (though the processes involved are still grossly oversimplified), it should be obvious that the most readily accessible element of this system, to those of us who are not telepathic, is the portion *outside* of the dotted box in the figure above. There are basically three aspects of this material: (1) the acoustic signal, (2) the physical structures and processes of the auditory system, and (3) the physical structures and processes of the articulatory system. It is the field of **phonetics** which investigates these matters, providing linguists with tools which allow them to get at the most accessible aspects of the speech system in order to test hypotheses about the underlying cognitive systems which play a key role in this system. In keeping with Figure 2.1, there are three major branches of phonetic science.

- The major branches of phonetics
  - (a) acoustic phonetics
  - (b) articulatory phonetics
  - (c) auditory (or perceptual) phonetics

In this chapter, we will be focusing on the study of *articulatory* phonetics. The reasons for this are simple: the study of acoustic and auditory phonetics generally requires access to sophisticated equipment, laboratories, anechoic chambers, and so forth. By contrast, the physical equipment necessary for the study of articulatory phonetics is with us all the time.

## 2.2 Some further implications of Figure 2.1

While talking about the line between the primary interest of linguists (the mental system called 'the grammar' in Figure 2.1) and the systems which turn the output of the grammar into something we can observe the effects of (the acoustic signal), there are a few important related issues we should deal with.

It is not unusual to hear adult speakers produce strings of speech such as the following (taken from a recorded conversation contained in a Carterette & Jones, *Informal Speech*, University of California Press, 1974):[3]

- aɪ ʤəs rili ɪt ɪts ðə bɪgəst θrɪl
- I just really it it's the biggest thrill

[3]This is presented first in the alphabet of the International Phonetic Association, which you will learn about later in this chapter, then in regular English orthography. After you learn about the IPA, you should take another look at this passage and try to read it out loud.

This speaker clearly started out saying one sentence, then shifted midstream to another (without finishing the first). She also said the subject of the second sentence ('it') twice. The result is that she has said something which is *not* a sentence at all – it cannot be parsed by a grammar. It is difficult to interpret exactly what was intended (what was she saying about herself with 'I just really', for example?). Part of what she said ('it's the biggest thrill') is a perfectly good English sentence, and is the only part of the string of noise she produced that can be coherently interpreted as a sentence. Notice, for example, that although the string starts out as if it is saying something about the speaker ('I'), you cannot respond to this sentence by saying 'no, you don't' (or 'no, you aren't').

This can be made clearer if we imagine trying to teach someone 'English.' One seemingly sensible way to do this would be to give them sentences said by speakers of English. Speakers say things like 'it it's the biggest thrill' every once in a while. But this is not due to the *grammar* which is present in their mind – it is instead an effect of the imperfect physical system through which the grammar is forced to play its output. If we taught our learner to say 'it it's the biggest thrill' (as speakers of English do every once in a while), when *the learner's* body suffered the same electro-chemical glitch that caused our speaker to say 'it it's the biggest thrill,' our learner would say 'it it it's the biggest thrill' (mistakenly saying sentence-initial 'it' twice). To get our learner to say 'it it's the biggest thrill' with the same frequency as our speaker does, we have to teach them to say 'it's the biggest thrill.' If we do that, the learner's *body* will produce – through the same processing imperfections that our speaker's body has – 'it it's the biggest thrill' every now and then (not that that is necessarily a desired outcome . . . ).

If 'it's the biggest thrill' is the real output of the grammar, we would expect our speaker to be able to tell us that. That is, even though speakers occasionally say 'it it's the biggest thrill,' they actually *intend* on those occasions to say 'it's the biggest thrill.' And, indeed, speakers do readily recognize that they occasionally fail to articulate precisely what they intended.

We therefore need to carefully distinguish between the **competence** of the speaker, responible for the output of their grammar (a mental representation), and their **performance** on some particular occasion (responsible for the acoustic output of their body). Bodily lapses of various sorts (lapses of memory, sneezes, etc.) and/or external physical disturbances (high winds, dogs jumping on your chest, and so forth) modify the acoustic signal which emerges from the body, but theses lapses/disturbances do *not* reflect parallel lapses/disturbances in the processes of the grammar. Lewis Carroll was toying with this distinction when he wrote, in *Through the Looking Glass*:

> 'And you do Addition?' the White Queen asked. 'What's one and one and one and one and one and one and one and one and one and one?'
>
> 'I don't know,' said Alice. 'I lost count.'
>
> 'She can't do Addition,' the Red Queen interrupted.

Of course Alice *knows* how to add a sequence of ones. But presented in rapid fire succession, her perceptual performance system prevents her arithmetical system from being able to access the information necessary to answer the White Queen's question. The perceptual and articulatory processing systems thus impose limits on access to underlying competence.

Consequently, although the output of the body provides linguists with much of the evidence they have for the underlying grammatical system, this evidence must be used with caution. In particular, it must be augmented by evidence which can be obtained by checking the judgement of the speaker as to whether or not a given string they produced involved a performance error. This evidence is known as a 'grammaticality judgement' – it can be used to help determine whether a string accurately reflects the output of the grammar or not. In general, for your own grammar, you can perform this empirical research through introspection.

## 2.3 Articulatory Phonetics

In order to produce sound, air needs to be set in motion. This is a fact about the world. Humans can make air move in a variety of ways, but in the case of speech only those methods involving the upper respiratory tract are used. The most widely used technique for moving air to create speech (called an 'airstream mechanism'), used in all human languages, involves *exhalation* (i.e., breathing out). It is known as the **pulmonic egressive airstream mechanism**. You can confirm the key role of the lungs in speech through the following simple experiment. Make a long (5-10 seconds) [s].[4] Try to pay

[4]It is traditional in linguistics to place phonetic representations of sounds in square brackets such as these.

close attention to what you are actually doing. Do this several times to get a feel for the component parts of the complex articulatory act. What's going on in your mouth? Now do it again but, in the middle of this long [s], stop the airflow coming up from your lungs without changing the position of the articulators in your mouth. All noise ceases. Without the flow of air from the lungs, no sound can be produced. Now make a long [s] again this time suddenly opening your mouth as wide as possible part way through it. The [s] sound will cease. This reveals the second critical aspect of producing speech: the modification of the flow of air by manipulating the path through which the air is flowing – in this case in the oral cavity (i.e., the mouth). (Airflow through this path with no modification does not make a very interesting noise.) Figure 2.2 shows, schematically, how these two components combine to produce acoustic output.

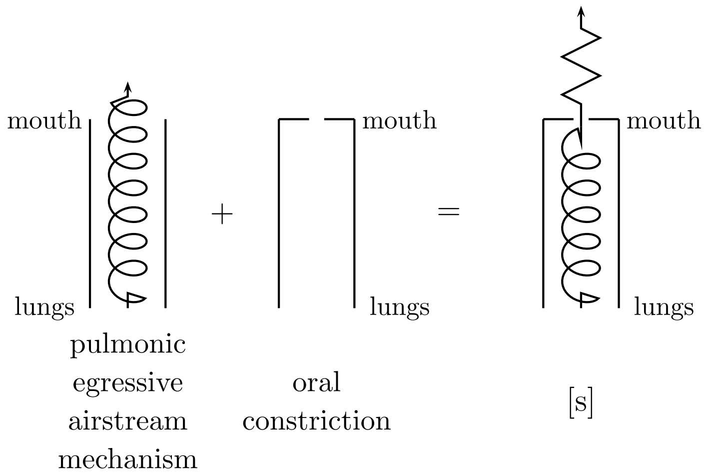



Figure 2.2: The components of articulation.

In Figure 2.2, one can see that the airstream mechanism alone (the leftmost schematic upper respiratory tract) will not produce an [s], nor will the oral constriction (in this case at the **alveolar** ridge, where [s]'s are made) do so alone. Only the *combination* of the two acts leads to the articulation of an [s]. In English-type grammars, all speech sounds are produced by superimposing a constriction in the vocal tract on exhalation. We will turn next to the question of what kinds of 'constriction' are involved.

## 2.3.1 Manners and Places of Consonantal Articulation

While the technical differences between consonants and vowels need not detain us in this introductory text, they consist basically of the fact that consonants involve greater 'constriction' of the airflow than do vowels. In general, it is more straightforward to describe the articulation of consonants, so we will begin there.

Produce the following: [apapapapapa].[5] Ignoring for the time being the vowels (to which we will return later), what are you doing to the flow of air when you pronounce the [p]'s? Produce the following: [ppppp] (that's a really long [p], no vowel). What's happening to the air in your lungs? Nothing. It is totally blocked. Sounds produced with a complete closure at

[5]The [a] indicates the vowel of the word 'father.'

some point in the vocal tract are called **stops**. Now produce the following: [atatatatata] and [ttttt]. Again, during the pronunciation of the consonants, the flow of air is totally blocked – stopped, as it were. [t] is therefore a stop as well. The various ways of constricting the airflow (e.g., stopping it completely) are referred to as **manners of articulation**.

If both [p] and [t] involve the total blockage of airflow during their articulation, how can we tell the difference between [apa] and [ata]? Why don't they sound the same? They sound different because the stop closure is being made *at a different point* in the vocal tract – when the air is released from the stop closure, the resulting disturbance (sound wave) is different for [p] than for [t].[6] Figure 2.3 is a schematic of the human head, with the major **places of articulation** in the vocal tract labeled. [See next page for figure.]

If the primary point of constriction for a consonant is at the lips (as for [p]), it is called a **labial**. If the primary point of constriction is against the back of the teeth, the consonant is called **dental**.[7] Directly behind your teeth, on the roof of your mouth, you'll find a hard little ridge, the 'alveolar ridge', which is the beginning of the so-called 'hard palate'. It is at this point that most speakers of English-type grammars in North American (and Southern England) produce, with the tip (or apex) of their tongues, their [t]-type sounds. They are called **alveolars**. Segments produced behind the alveolar ridge, on the hard palate, are known as palatals. At the end of the hard palate, the tissue becomes soft. The soft area is known as the velum. Segments produced in this area – generally by bringing up the back part of the tongue (the dorsum) – are called **velars**. Finally, hanging down at the end of the velum is the **uvula**. This is the dangling piece of soft tissue that is featured prominently when they show babies crying in cartoons. Sounds produced at this point in the oral cavity are called **uvulars**. The *glottal* point of articulation will be explained in some detail below.

[6] This is a simplified picture of the actual acoustic description.

[7] Most Western European-type grammars (but not English-type ones) pronounce sounds like [t] in the dental point of articulation.

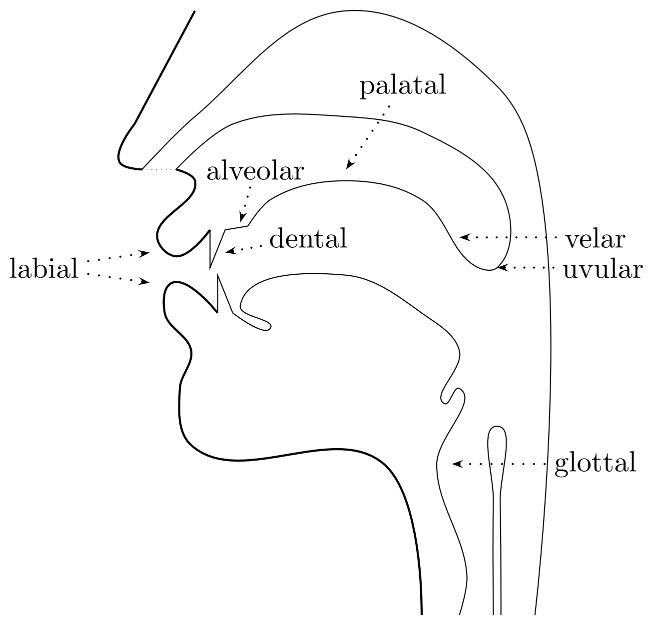



Figure 2.3: Places of articulation.

Consonants are normally described using the names of their place and manner of articulation. [p] is thus a *labial stop*, while [t] is, in most English-type grammars of North America, an *alveolar stop*. If you say [akakakakaka] and [kkkkk] you will realize that [k] is produced further back in the vocal tract than either [p] or [t]. [k] is a *velar stop*. (English-type grammars do not have palatal stops, which are rare, cross-linguistically, although they do exist.)

The International Phonetic Association developed an alphabetic system to represent the sounds of natural language. This was motivated by the need to describe and represent sounds consistently, independent of the writing system (or lack of writing system) of any particular language. The International Phonetic Alphabet (IPA) is supplemented by a set of diacritic marks (mainly superscripted and subscripted notations) to increase its utility. While it is true that there is a certain amount of intra- and inter-speaker variation in the utterance of any one particular sound (for example, every time you produce a [p], the actual acoustic result is slightly different), the IPA captures to a sufficient extent all the critical differences in natural language sounds. You can easily see the necessity for such a system if you compare the following words in English orthography (traditionally, this terms refers to the spelling system but is most often now used for a writing system) – ‘rough’ and ‘though’. The ‘gh’ in ‘rough’ represents a final consonant sound of [f] in ‘rough’ and represents *no* sound in ‘though’ (which ends in the vowel sound [o]). While English is a rather extreme case of inconsistent symbol-sound representation, all writing systems have this problem to some extent. The IPA allows us to overcome these difficulties.

For the time being, we will introduce only those IPA symbols which are used to represent sounds in English. Further symbols will be added later as needed. Many of the symbols are common to the English writing system. For example, the IPA symbol for the velar stop is [k], the IPA symbol for the alveolar stop is [t], and the IPA symbol for the labial stop is [p]. The task of remembering these is trivial, however, using them correctly takes a certain amount of attention since the symbols represent *sounds*. Thus, the initial sound of <cat> is represented by [k], the initial sound of <pterodactyl> is represented by [t], and the initial sound of <pneumonia> is *not* represented by [p] but rather by [n].[8]

English has a stop in yet another place of articulation. This one really does have a special symbol which must be learned. In addition, we do not typically write this stop at all in standard English orthography so it may be surprising that such a thing exists. For most North Americans, <mountain> is pronounced as [maʊʔn̩]. ‘He hit me’ is pronounced as [hi hɪʔ mi]. The sound indicated by [ʔ] in the IPA is a **glottal stop**.

To understand how glottal stops are produced, as well as a number of other sounds, it is necessary to describe the **glottis**. Midway down your neck, noticable (especially on males) as a bony protrusion (called an ‘Adam’s apple’), is your larynx. It is sometimes called a ‘voice-box.’ Inside your larynx are folds of tissue known generally as the ‘vocal chords.’ The muscles in the larynx allow manipulation of these vocal folds, controlling the amount of tension in the folds themselves. Normally, when you’re just breathing, these folds are apart and air flows relatively freely between them. However, you can pull the vocal folds closed with varying degrees of strength. The space between the folds is called the **glottis**. When pulled tightly closed, so tightly that the air coming up from the lungs can’t get through the vocal folds, the glottis is closed and you are making a glottal stop. Practice doing this by

[8] Carat brackets (¡ ¿)are used to indicate orthographic representations when it is necessary to distinguish these from phonetic representations.

saying [aʔaʔaʔaʔaʔa]. (If the indications above didn't make it clear to you exactly what a glottal stop is, perhaps this will help: it is the stop in the middle of the widely-used negator [əʔə], sometimes written 'uh-uh.')

This does not exhaust the stops of most varieties of English. If you say [abababababa] you will notice that there is a blockage of the airflow in that sequence as well. Where is the stop closure being made? It is labial, obviously, but we already have a labial stop, [p]. If you compare [apa] to [aba] you will notice that, in both cases, your lips are temporarily pressed together. Why, then, do these segments sound different from one another? The answer lies in the different states of the glottis.

In addition to being tightly closed (as for a glottal stop) or wide open (for breathing), the glottis can adopt another posture. If you pull your vocal cords loosely closed and then push air through them (by normal exhalation), they will vibrate. You can mimic this behavior with your lips. If you hold them very tightly closed, they will stop the flow of air. If you hold them apart, the air will flow between them unimpeded. But if you hold your lips loosely together and then push air through them, you will get a mildly embarrassing kind of noise from the flapping of your lips in the passing breeze. When you make this kind of noise in your glottis, the resonating cavity of your head shapes it into that mellifluous sound which is your voice, much as the lip-buzzing sound that a trumpet-player blows into the business end of a trumpet is converted by the resonating cavities in the trumpet into pleasant trumpety sounds.

The technical name for the sound produced by this rapid flapping of the vocal folds is **voicing**. Segments which are produced while the vocal folds are vibrating are therefore called 'voiced' sounds. Since this voicing takes place some distance away in the glottis, it can be superimposed upon oral constrictions of various types, e.g., stops. The difference between [p] and [b], both of which are labial stops, is that during the production of [p] the vocal folds are pulled away from one another, allowing the air to flow through them unobstructed (although the air is ultimately blocked by the lips). During the production of a [b], by contrast, the vocal folds are pulled together, vibrating in the flow of air coming up from the lungs. [b] is therefore a *voiced labial stop*, while [p] is a *voiceless labial stop*.

You can establish for yourself that this is the case with a simple experiment. Note that you can make a very long [p] – in fact, you can make a [p] that lasts just as long as you can hold your breath. On the other hand, if you try to make an equally long [b] you will discover that you have a little aerodynamic problem. To keep the vocal folds vibrating (which you must do so that the [b] does not turn into a [p]), you have to keep pushing air up from your lungs through the loosely-closed glottis. But your lips are closed, so there's nowhere for the air coming up from lungs to go. Once the capacity of your oral cavity is exhausted, you can't push any more air through the glottis and voicing will cease. (Attempting to exceed the capacity of your oral cavity will cause the look and feel of imminent explosion.)

There are voiced versions of the alveolar and velar stops as well. [d] as in <duck> is a *voiced alveolar stop* and [g] as in <get> is a *voiced velar stop*. There is, of course, no voiced glottal stop. Since to make a glottal stop you must hold your vocal folds tightly closed, allowing no air to escape, and to produce voicing you have to push air through the glottis causing the vocal folds to vibrate, it is physically impossible to do both simultaneously.

There is yet another sound made with the lips blocking the flow of pulmonic egressive air. This sound is [m] as in <mouth>. Can you tell whether [m] is voiceless or voiced? Assuming that you replied 'voiced', you are correct. You can feel your vocal chords vibrating during the articulation of [m] if you press your fingers to your throat at the point of your larynx and say [mmmmm ðæts gʊd] ('mmmmm that's good'). We pointed out above that you get an 'explosive' feeling if you try to make a very long [b] because of a build-up of pressure from the need to push air through the glottis constantly to keep up the voicing. But if [m] is voiced, why doesn't your head explode when you make a really long [m]? (You may safely test this.) Obviously, the pressure is being released somehow – the air is getting out. Just as obviously, the air is not being released through your mouth, which is closed at the lips. This means it must have another escape route. Given the anatomical structure of the human head, the possible escape routes are limited. A good candidate would be one of those orifices connected to the vocal tract, since that's where the pressure is building up in the case of a long [b]. Now, as you know if you have ever (1) burst out laughing while taking a drink, or (2) unwisely left your mouth open when diving into water, your *nasal sinus* passage is directly connected to the oral cavity.[9] The relationship between the nasal and oral cavities can be seen in Figure 2.3 where the velum is where air will divert between the two passages.

[9] It is true that the ears, and even the eye sockets (in some mysterious way), are also connected to the naso-oral cavity, as demonstrated several times by invitees to a certain late-night TV show. These connections are a lot less direct, however, and air is not released through these passages during speech.

From Figure 2.3, it would appear that air should always be able to pass through the nasal cavity, but this is not the case. The velum is a 'movable' part although it is difficult to sense this movement consciously. In particular, it can (and does) move back and block access to the nasal passage from the oral cavity. Fewer nerve endings in that general area reduces your ability to feel the velaric movement, but you may try the following. The vowel written [ɛ] in the IPA is the vowel of <bet> [bɛt]. The vowel written [ɛ̃] is the vowel of French <fin> ('end')[fɛ̃].[10] Pronounce the following: [ɛɛ̃ɛɛ̃ɛɛ̃ɛɛ̃ɛɛ̃]. It's pretty hard to do that, actually, except very slowly. If you do it, you will notice that nothing in your mouth moves at all, nor does the airflow through the glottis vary (both [ɛ] and [ɛ̃] are voiced). Why do the two vowels sound different? Because the velum is moving back & forth to open (for [ɛ̃]) and close (for [ɛ]) access to the nasal sinuses.

The class of consonants produced with the velum open (i.e., lowered) are called **nasals**. As we have already established, [m] is a labial nasal. Parallel to the alveolars [t] and [d] we have the alveolar nasal [n], as in <nasal> [nezl]. Parallel to the velar stops [k] and [g], we have the velar nasal [ŋ] as in <sing> [sɪŋ]. Nasals are typically classified as stops, since there is an oral occlusion at the normal stop positions.

To summarize, the set of stops (including the nasals) is given in the table in Figure 2.4.

| | labial | alveolar | velar | glottal |
|---|---|---|---|---|
| voiceless | p | t | k | ʔ |
| voiced | b | d | g | |
| nasal (voiced) | m | n | ŋ | |

Figure 2.4: Stops.

The official way to refer to stops, and consonants generally, is to state their voicing, place, and manner, in that order. [b] is thus a 'voiced labial stop,' [n] is a 'voiced alveolar nasal'[11] and [k] is a 'voiceless velar stop. Every speech

[10] If this is not a promising clue, it is also close to the sound that Bugs Bunny made at the beginning of his standard question 'What's up doc?' usually written as 'eh' and articulated as though he had a cold.

[11] Since nasals are almost always voiced, the 'voiced' designation is sometimes left off for nasals.

sound must, in addition, be either oral or nasal. (The velum will either be raised or lowered.) Because the default value for speech sounds appears to be 'oral' (more sounds are oral sounds than are nasal sounds), we typically specify this parameter only for those sounds which are nasal (i.e., we do not usually call [k] a voiceless velar **oral** stop, for example, although it would be perfectly correct to do so).

The segments we have examined so far all involved total occlusion of the airflow through the oral cavity; however, it is possible to disrupt, rather than completely block, the flow of air. One technique for doing this is to force the air flow to go through a very narrow passage. Since the air is coming up from the lungs at a more or less constant rate, when it is forced through a small opening it must speed up significantly to get out of the way of the rest of the air rushing up from the lungs. The narrow channel causes this rapidly rushing air to become quite turbulent and these sounds are characterized by a great deal of 'noise' (in the technical sense). Sounds that are made in this manner are called **fricatives**. A typical example of a fricative, quite widely found in the languages of the world, is the voiceless alveolar fricative [s] (as in 'Simpsons' [sɪmp**s**n̩z]). An [s] is articulated by using the apex (tip) of the tongue to block all air flow through the oral cavity except for a narrow channel between the tongue and the alveolar ridge. [s] has a voiced counterpart in [z] (also in 'Simpsons,' but at the end). Indeed, a very good way to get a feel for exactly what is going on with your vocal chords is to pronounce [szszszszsz] – when you do this, you can feel the vocal chords engaging as you go from [s] to [z].

Grammars of the English type, like many grammars around the world, lack labial fricatives, though they do exist in some languages. Acoustically very similar to the labial fricative is the **labiodental** fricative, which is produced by forcing the air through a very small set of channels between the upper teeth and the lower lip. The official IPA symbol for the voiceless labiodental fricative is [f], as in 'failure' [fæljr̩]. The voiced labiodental fricative is [v] as in 'veggie' [vɛʤi].

Most English-type grammars also have a pair of **interdental** fricatives – quite rare, cross-linguistically. These sounds are produced by placing the apex of the tongue between the teeth, forcing the air through small channels between the teeth and the tongue. The voiceless interdental fricative is found in words like 'thrilling' [θrɪlɪŋ]. The voiced interdental fricate is found in words like 'bathe' [beð].

The **alveopalatal** fricatives are produced with the blade of the tongue making a narrow closure in the region just behind the alveolar ridge at the very frontmost edge of the (hard) palate. There is a voiceless one, [ʃ], as in 'shake' [ʃek], and a voiced one, [ʒ], as in 'measure' [mɛʒr̩]. Frequently, a non-standard (i.e., non-IPA) symbol is used for these sounds in North America. The voiceless alveopalatal fricative is often written [š] and the voiced one [ž]. We will continue to use the standard, IPA symbols for these sounds but they may appear in their non-standard forms elsewhere.

English type grammars typically lack both palatal and velar fricatives, though such segments are not rare cross-linguistically. Many linguists treat [h], as in 'hop' [hɔp], as a fricative. In this case, the frication/turbulence is found at the glottis (caused by air rushing through vocal folds that are drawn together slightly). We will treat [h] as a glottal fricative here.

It is useful to deal with a rather peculiar type of segment at this point – one called an **affricate**. The affricates start out with a stop closure, but instead of simply releasing the closure and allowing the air to flow freely after the stop release, only a small portion of the closure is released, leading to the type of narrow channel generally characteristic of fricatives. These segments are thus combinations of stops and fricatives, neither having the full duration of corresponding full segments of those types. English has two of these affricates, a voiced-voiceless pair. They are produced in the alveolar and alveo-palatal area (the stop portion being essentially alveolar, though slightly further back, and the fricative portion being essentially alveo-palatal). They are called the alveo-palatal affricates, the voiceless [ʧ], as in 'chunk' [ʧəŋk], and the voiced alveo-palatal affricate, [ʤ], as in 'junk' [ʤəŋk]. As with the alveo-palatal fricatives, there are special symbols in widespread use in North America for these two segments. The voiceless alveo-palatal affricate is frequently written [č] and the voiced one [ǰ].

To summmarize what we have covered about fricatives and affricates, we have included the chart in Figure 2.5.

| | labio-dental | interdental | alveolar | alveo-palatal | glottal |
|---|---|---|---|---|---|
| Fricatives | | | | | |
| voiceless | f | θ | s | ʃ (=š) | h |
| voiced | v | ð | z | ʒ (=ž) | |
| Affricates | | | | | |
| voiceless | | | | ʧ (=č) | |
| voiced | | | | ʤ (=ǰ) | |

Figure 2.5: Fricatives and affricates.

We see the same general pairing of voiceless and voiced segments at each place of articulation (except for glottal, where it would make no sense) as we saw in the stop table in Figure 2.4.

The remaining segments are harder to describe the articulation of and they do not, unfortunately, fit into nice neat tables. English-type grammars have two **liquids**: [r] and [l] as in 'really' [rili]. The articulation of [r] is quite complex. It is formally characterized as a 'retroflex approximant' – 'retroflex' because the tongue tip is curled backward (such that the underside of the tip is parallel to the alveolar ridge) and 'approximant' because the tongue *approaches* but does not touch the alveolar ridge. The symbol [r] is not the standard IPA symbol for this sound – the IPA symbol is [ɹ]. Due to the difficulty of accurately reproducing this symbol when writing by hand as well as the difficulty in remembering the exact orientation of it, we will not use it here but will substitute orthographic [r]. (The latter is also an IPA symbol – it stands for the sound of a 'trill'.) Of the many r-type sounds found in natural language, the English type is most unusual. The [l] is also, technically, an approximant, but a **lateral** approximant. It is made by pushing the apex of the tongue up against the alveolar ridge and allowing the air to flow around the sides of the tongue. The 'lateral' designation is because of this pattern of airflow. While it may not, at first glance, seem to fit into the approximant category, as described above, the critical portion of the articulation of [l] is its 'laterality' – the areas where the tongue *approximates* but does not actually touch, the sides of the alveolar ridge. For convenience, we will continue to use the term 'liquids' for [r] and [l]. To distinguish between these two liquids, the terms *rhotic* (after the Greek *rho* and *lateral* (as above) are used for r-like and l-like sounds, respectively.

Nasals and liquids, in English, share something in common with vowels.

They may form what is called the *nucleus* of a syllable all on their own. The nucleus might be described as the 'heart' of the syllable – if there is no nucleus, there is no syllable. Interestingly, everyone has good intuitions about syllables and this really means that they have good intuitions about syllabic nuclei. For example, if we were to say aloud a long and unfamiliar word, you would be immediately able to tell us how many syllables were contained in it. People appear to do this in the same way that they can follow the beat of music by tapping their fingers or feet. For example, the word 'antidisestablishmentarianism' has twelve nuclei and, thus, twelve syllables. Usually, it is vowel sounds which carry the 'beat' (i.e., form the nuclei) however liquids and nasals may do so as well under certain conditions. When a consonant, such as a liquid or nasal, acts as a syllabic nuclei, we call it a *syllabic* consonant. Syllabic status for liquids and nasals is indicated by a little bar under the standard IPA symbol for that segment as in the following examples:[12] 'master' [mæstr̩], 'bird' [br̩d]; 'snap, crackle and pop' [snæp krækl̩ n̩ pap], 'castle' [kæsl̩]; 'button' [bəʔn̩], 'fasten' [fæsn̩]; 'something' [səʔm̩], 'chasm' [kæsm̩].[13] Certain aspects of syllable structure are complex and still not very well understood. The brief notes above will be sufficient for our purposes.

As far as the major consonants needed to discuss most English-type grammars in a somewhat superficial way, there are only two segments left. These are usually called *glides* or *semivowels*.[14] The *labiovelar* glide is transliterated [w] and found in words such as 'wacky' [wæki]. It is made by rounding the lips and retracting the back of the tongue toward the velum. The *palatal* glide has the official IPA symbol [j] and is found in words such as 'yucky' [jəki].[15] In this case, the blade of the tongue is placed very close to the palate but not close enough to cause frication. The alternate 'semivowel' designation is the result of the fact that these sounds are virtually identical to two of the vowel sounds [u] and [i] (discussed below in Section 2.3.3) except in their position within the syllable (critically *not* the nucleus position). Their

[12] This subscripted bar is one of the set of diacritics used to supplement the IPA that we mentioned earlier.

[13] Some people, transliterate the syllabic [r] with a vowel symbol specially designed for that purpose, [ɚ].

[14] Strictly speaking, phonetically, these two sounds also fall under the 'approximant' category but it is more common and, we think, less confusing, to use the term 'glide'.

[15] Many people use the symbol [y] to designate this sound, but as the IPA symbol [y] is officially used for a relatively critical vowel sound that exists in a number of languages, it is a good idea to keep these distinct.

position within the syllable is like that of consonants.

### 2.3.2 Some Further Consonantal Details

Although not always covered in introductory texts, we consider the following sounds, universally present in English-type grammars, to be worth describing. First, there are dental stops in addition to labial, alveolar, velar and glottal. If you contrast the pronouncation of 'ten' [tɛn] and that of 'tenth,' you will find that the latter should be transliterated [tɛn̪θ], where [n̪] is the symbol for the *dental nasal*. Similarly, for those speakers who have a stop before the final fricative in a word such as 'eighth', it is a *voiceless dental stop*: [et̪θ]. More frequent is the *voiced alveolar flap* or *tap* (use either term). This sound is found in the middle of words like 'middle' and 'butter,' and its official IPA symbol is [ɾ]. 'Middle' and 'butter' should therefore be transliterated [mɪɾl̩] and [bəɾr̩], respectively.

### 2.3.3 Vowels

English-type grammars are extremely rich in vowels. Whereas with consonants there were only a few 'funny' symbols of the IPA that needed to be learned, in the case of vowels virtually *all* of the symbols have unexpected values for speakers of English. This may seem burdensome but since vowel symbols will be in extensive use, it is unavoidable.

Consonants are generally classified according to the point of greatest constriction in the vocal tract – the number of such points used in human language is quite restricted. Vowels, by contrast, are distinguished from one another by the shape of the (usually oral) resonating cavity. The shape of this cavity can be modified in three major ways. The lower jaw can be raised or lowered (making the resonating cavity smaller or larger, respectively). The body of the tongue can be bunched up in the back of the oral cavity or the blade of the tongue can be pushed up toward the palate. Finally, the lips may be rounded or spread during the articulation of the vowel. Thus vowels are primarily defined by the following parameters:[16]

- height of the lower jaw (high, mid, low)

[16] We will see below that additional features may distinguish vowels – nasalization, for example – but the most important for our purposes in this class are those given here.

- position of tongue (front, central, back)
- lip rounding (rounded, unrounded)

If you raise your lower jaw as high as possible (without making such a tight constriction in the oral cavity that you make a consonant) and push your tongue forward with your lips spread (i.e., not rounded) you can produce the sound [i] of 'beat' [bit]. [i] is therefore a high, front, unrounded vowel. When the doctor wants to look down your throat, he or she says 'say [a]' (if he or she says 'say [i]', it may be time to look for a new doctor) because [a] is a *low* vowel (low, central and unrounded). [a] is the vowel of 'father' [faðr̩]. The vowel [u], like [i] is made with the lower jaw up (i.e., it's a high vowel). However, unlike [i], to make an [u] the back of your tongue is bunched up towards the velum (but not close enough to cause frication) – it is therefore a back vowel. Like all back vowels in English-type grammars, [u] is round.

These three vowels more or less define the edges of the vowel space. You should practice making the following sequence of sounds: [iuaiuaiuaiuaiua]. This will help you get a feel for the space within which most vowels are produced. This space is typically represented as in the figure below.

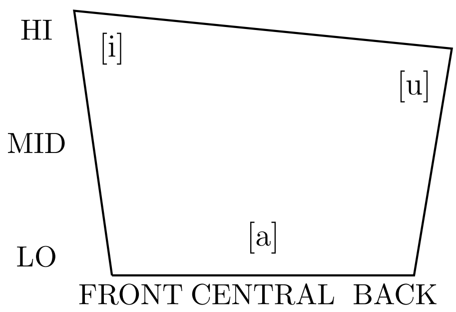



Figure 2.6: The vowel space.

English-type grammars have many more vowels than these three, of course. For example, if you say [iaiaia] very slowly and stop the jaw-lowering involved about halfway between the end of the [i] and the start of the [a] you'll be producing a mid, front, unrounded vowel: [e], as in [bet] 'bait.' Similarly, in the back area, if you make a vowel about half-way between a high [u] and a low [a], you'll be producing an [o], as in [bot] 'boat.' These five vowels are very widely found in the languages of the world.

At the high and mid heights, however, English-type grammars have additional vowels. These contrast with the vowels cited above at the high and mid points by virtue of the degree of vocal tract constriction required for their articulation. Vowels such as [i] and [u] are generally called 'tense' vowels because the muscles of the vocal tract are constricted during their articulation. By contrast, the vowel of a word like 'bit,' which is, like [i], high, front and unround, is called 'lax'. The IPA symbol for the vowel of 'bit' is [ɪ], thus [bɪt]. Corresponding to tense [u] in the high, back, round space we have the lax vowel [ʊ] of 'put' [pʊt]. Articulated in the mid, front, unround area (near [e]) is the lax vowel [ɛ] of 'bet' [bɛt]. Finally, corresponding to the tense version of the mid, back, round vowel ([o]), we have a lax version [ɔ] in words such as 'law' [lɔ].

One of the big differences between various dialects of English concerns the distribution of vowel sounds in specific words, often the sounds [ɔ] and [a]. In some dialects, 'caught' is pronounced as [kɔt] and 'cot' as [kat]. In others, the two words are homophonous – both are pronounced as [kɔt]. Similarly, some speakers say [pat] and [sari] for what others pronounce as [pɔt] 'pot' and [sɔri] 'sorry.' You can see how convenient the IPA is for allowing us to describe the differences between various linguistic systems!

English-type grammars have two other simple vowels. The first is the low, front, unround vowel [æ] of [bæt] 'bat.' The other is the so-called *schwa* (its name is pronounced [ʃwa]), whose symbol is [ə]. This vowel is found in words such as [bət] 'but' or [sofə] 'sofa.' Both [æ] and [ə] are lax, as is the other low vowel of English [a].

The set of simple vowels which we've introduced for the discussion of English-type grammars is given on an articulatory chart in Figure 2.7.

However, we are not yet done with the vowels. In addition to these 'simple' vowels, English has a number of *diphthongs* – i.e., vowels which involve movement from one vowel position to another during their articulation. By tradition, they are transliterated as sequences of vowel+glide. These diphthongs include [ɔj] as in [bɔj] 'boy,' [aj] as [bajd] 'bide,' and [aw] as in [bawt] 'bout.' The trajectory of tongue/jaw movement during these diphthongs is indicated in Figure 2.8.[17]

[17]To be technically precise, the tense mid vowels of English are also diphthongs, the [e] raising up towards the [i] space during its articulation, the [o] raising towards [u]. This

## 2.4 Summary of relevant IPA symbols

The tables below present listings of the symbols we have studied so far as well as some examples of words which generally, in English-type grammars, exemplify the use of these symbols. Again, the critical descriptive information for consonants consists of place of articulation, manner of articulation, and voicing. For the vowels, what is important are the features height, backness, lip-rounding, and tense or lax.

| | labial & labio-dental | dental & inter-dental | alveolar | alveo-palatal | palatal | velar | glottal |
|---|---|---|---|---|---|---|---|
| *stops* | | | | | | | |
| voiceless | p | t̪ | t | | | k | ʔ |
| voiced | b | d̪ | d | | | g | |
| nasal | m | n̪ | n | | | ŋ | |
| *fricatives* | | | | | | | |
| voiceless | f | θ | s | ʃ | | | h |
| voiced | v | ð | z | ʒ | | | |
| *affricates* | | | | | | | |
| voiceless | | | | ʧ | | | |
| voiced | | | | ʤ | | | |
| *liquids* | | | | | | | |
| rhotic | | | r ɾ | | | | |
| lateral | | | l | | | | |
| *glides* | w | | | | j | | |

Figure 2.9: The primary consonants of English.

causes some people to transliterate them [ej] and [ow]. We do not do this, for reasons which are too complicated to go into.

| | |
|---|---|
| ʔ | [bəʔn̩] ‘button’; [hi hɪʔ mi] ‘he hit me’ |
| θ | [θrot] ‘throat’; [bæθ] ‘bath’; [tiθ] ‘teeth’ |
| ð | [ðiz] ‘these’; [beð] ‘bathe’; [tið] ‘teethe’ |
| ʃ | [liʃ] ‘leash’; [mɛʃ] ‘mesh’; [ʃɪvr̩] ‘shiver’ |
| ʒ | [liʒr̩] ‘leisure’; [mɛʒr̩] ‘measure’ |
| ʧ | [ʧr̩ʧ] ‘church’; [bæʧ] ‘batch’ |
| ʤ | [ʤəʤ] ‘judge’; [bæʤ] ‘badge’ |
| ɾ | [bəɾr̩] ‘butter’; [mæɾr̩] ‘matter’ & ‘madder’ |
| j | [jɛlo] ‘yellow’; [kjut] ‘cute’; [traj] ‘try’ |
| i | [bit] ‘beat’ and [fit] ‘feet’ |
| ɪ | [bɪt] ‘bit’ and [fɪt] ‘fit’ |
| e | [bet] ‘bait’ and [fet] ‘fate’ |
| ɛ | [bɛt] ‘bet’ and [fɛt] ‘fête’ |
| æ | [bæt] ‘bat’ and [fæt] ‘fat’ |
| u | [but] ‘boot’ and [fud] ‘food’ |
| ʊ | [pʊt] ‘put’ and [fʊt] ‘foot’ |
| o | [bot] ‘boat’ and [flot] ‘float’ |
| ɔ | [bɔt] ‘bought’ and [lɔ] ‘law’ |
| a | [faðr̩] ‘father’ |
| ə | [bət] ‘but’ and [(ɛlmr̩) fəd] ‘(Elmer) Fudd’ |
| aj | [bajd] ‘bide’ and [faj] ‘fie’ |
| ɔj | [bɔjd] ‘Boyd’ and [fɔjl] ‘foil’ |
| aw | [bawt] ‘bout’ and [lawd] ‘loud’ |

Figure 2.11: Examples for ‘unusual’ IPA symbols.

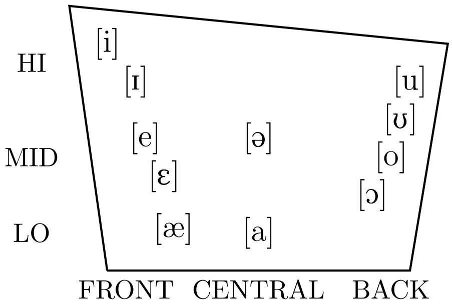



Figure 2.7: The simple vowels of English.

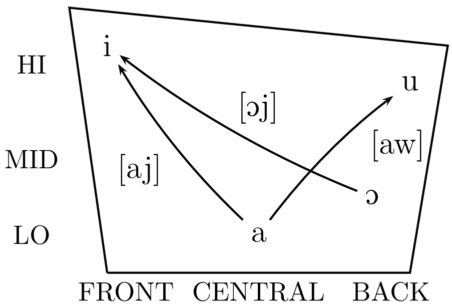



Figure 2.8: The diphthongs of English.

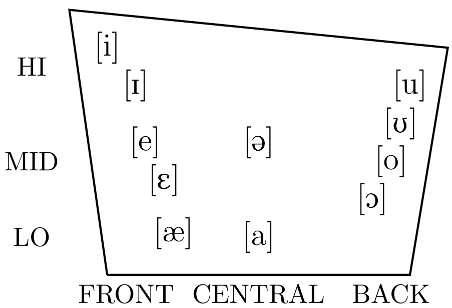



Figure 2.10: The simple vowels of English.

A number of the 'unusual' IPA symbols have their own names. Since the names frequently have something to do with the sounds represented by the symbols, they may be helpful to learn. Greek symbols retain their Greek letter names: θ is 'theta'; ɛ is 'epsilon'; and ʊ is 'upsilon.' Two symbols are borrowed from Old English: ð is 'eth' (pronounced [ɛð]; and æ is 'ash.' For ease of reference alone, we include the following: ɔ is called 'open o' and the ə is called 'schwa.'

## 2.5 Exercises

1. [θ] is a

   (a) voiced labiodental fricative.

   (b) voiceless alveolar fricative.

   (c) voiceless interdental fricative.

   (d) voiceled palatal fricative.

2. A voiceless alveolar fricative is transcribed in the IPA with the symbol

   (a) [ʃ]

   (b) [s]

   (c) [θ]

   (d) [⊙]

3. [u] is a

   (a) a high, back, round vowel.

   (b) a high, front, unround vowel.

   (c) a high, back, unround vowel.

   (d) a high, front, round vowel.

4. Examine the following dialogue between children and answer the question. [Note: This and the following excerpts were transcribed from an actual conversation between young children. They are mildly incoherent and have many performance errors thus they are quite similar to adult speech . . . ]

A: aj æ fɔr sɪstr̩z.

B: ju dən hæv fɔr sɪstr̩z. jə oni hæv θri.

A: suzi kæθi mi n̩ bani.

B: ʔɪts na kawntɪŋ ju. jr̩ naɾ ə sɪstr̩. jr̩ naʔ jr̩ sɪstr̩.

How many sisters does Miss A have?

(a) Four, but she thinks she has three.

(b) Three, but she thinks she has four.

(c) Five, but she thinks she has 4.

(d) The details of Miss A's family situation cannot be deduced from this dialogue.

5. Examine the following child's monologue and answer the question.

A: aj hæv ðiz awʔ ɔrɛɾi ɪn ðɪs wən kem awt. læs najt maj mawθ əz ɔl bliɾɪŋ. no, naʔ frəm ðæt, maj hol ɡəm ərawn æn maj hol tiθ. ɪt ɔmos ɪʔ drɪpt awɾ ə maj mawθ. maj brəðr̩ hæʔ wɛl hi hæd e tiθ awʔ. hɪz mawθ wəz ɔl bliɾi ən i hæʔ hi kunnivən stæn frəm ɪt. ɪʔ ɪɾɔlmos drɪpdawn tu ɪz to. tu ɪs fit. æn maj mam hæɾ ən apr̩eʃɪn. ʔapr̩eʃɪn ɔnr̩stəmɪk.

How far did the blood from A's brother's mouth almost drip down to?

(a) To his bellybutton.

(b) To his knees.

(c) To his ankles.

(d) To his toes or to his feet, the text is a little unclear on the point.

6. Examine the following children's conversation and answer the question.

A: pɛpr̩z ə ʤr̩mɪn ʃɛpr̩d.

B: jə no wət. ʔə bɔj sɛd ðɛt hi ðɛt pɛpr̩ hɛd ʔəm rebiz. bət mi n̩ hɪm wən ovr̩ tə pɛɾ ɪt. ɛn ɪ dɪdn̩ ivn̩ æv rebiz. hi wəz rɔŋ.

A: pɛpr̩ dɪd.

B: aj pɛɾəd hɪm. hi dɪdn̩ ivn̩ hæv rebiz. f i hæd rebiz hi wʊɾə bɪʔ mi. bɪʔn̩ mi ʔn̩ ʔəm dajd. əhə.

C: maj dɔgz hævɪŋ bebiz n̩ hi dɪdn̩ bajt.

B: hæv rebiz.

C: bebiz.

B: bebiz. wɛl ðe donʔ bajʔ wən ðe hæv rebiz. bebiz. bət. wən ðe ʔæv ʔm̩ rebiz jæ ðe du bajt.

Which of the following statements is true?

(a) According to A, 'Pepper' had rabies.

(b) According to B, 'Pepper' did not have rabies.

(c) C is a little confused about whether the discussion concerns 'babies' or 'rabies.'

(d) All of the above.

7. Examine the following child dialogue and answer the question.

A: wəns wəns hi kip ʤəmp ʤəmpɪŋ ʤəmpɪŋ æn ðɛn ðɪs θɪŋ slajs awn ðɛn hi fɛl ovr̩ tu ðəʔ əðə bɛd æn i star krajɪn ɛn aj kʊʔən gɛ tə bɛd so aj aj hæftə wek əp pəʔəm bæk ɪm maj krɪb.

B: ʔɪn jr̩ krɪb.

A: no naʔ ɪm maj krib. aj don hæv ə krɪb.

B: yu sɛ pəʔɪm bæk ɪn jr̩ krɪb.

A: aj min ɪn hɪz krɪb. aj don hæv ə krɪb.

Does Miss A have a crib?

(a) She said she did, but she doesn't.

(b) She said she didn't, but she does.

(c) No, B has a crib.

(d) No, B said she had a crib (but she doesn't).

8. Examine the following first-grader monologue and answer the question.

A: bɪfɔr aj juz tə hæv ə br̩d. ɪʔ dajd. aj juztə hæv ə ræbɪt. ðæʔ dajd. wi gaɾə dɔg ðæʔ dajd. ən wən wi gaɾə kæt wi sold ɪt əwe. so wi gat dɪsajdəd tə gɛɾ ənəðr̩ dɔg.

Given A's past experience with pets, what are the prospects for his new dog?

(a) Promising – it will probably lead a long and healthy life.

(b) Excellent – it can look forward to many years of happy pethood.

(c) Fine – I'm sure it will be well taken care of.

(d) Not so great, it should quickly find itself a new home.

9. Examine the following first-grader conversation and answer the question.

A: æn aj hæv ə θawzɛnz ɔmos θawzənz əv ænəml̩z ɪm maj bɛəd. aj oni ga wən tr̩dl̩.

B: aj ləv ræbɪs.

C: aj gaʔʔəbawʔ tɛn fajv ræbiʔs. fajv ræbɪʔs.

A: aj hæv ə dɔg. hɪz nem ɪz dæjəndi.

B: aj əv ə pɛwokiʔ bəri dəzən tɔk jɛt.

A: æn hiz ə æn izə winr̩ dɔk. hi wɪgl̩z ɪz teəl ɛvri tajm.

B: du ju kɔlɪm ɛnitajm ə winr̩. ɛnitajm.

A: hi ʤəmʔs ɔn ə ʧer̩z ən ɛvriθɪŋ. maj faðr̩ waps ɪm wən.

B: aj gaɾə slo ænɪml̩. ən gɛs hu hi ɪz. ðə tɔrdɪs.

What is B's slow animal?

(a) A rabbit.

(b) A wienerdog.

(c) A tortoise.

(d) A parakeet.

10. One of the four strings below is likely to be accurate – the others have serious errors which make it very unlikely that you or we would ever say them that way. Circle the one which has the good transliteration in it.

(a) ɪf ju kn̩ rid ðɪs, jə məjt pæs ðɪs klæs.

(b) ðiz kwɛʃʧn̩z ar rili stupid.

(c) maj hɛad hurts.

(d) ði aj pi e ɪz ə ril pain ɪn ði æs.

# Chapter 3

# Phonology

## Contents


3.1 Further Phonetic Details . . . . . . . . . . . . . . 46
3.2 The Phoneme . . . . . . . . . . . . . . . . . . . . . 48
3.3 Minimal Pairs & Allophones . . . . . . . . . . . . 50
3.4 More English allophony . . . . . . . . . . . . . . . 53
3.5 Rule ordering . . . . . . . . . . . . . . . . . . . . . 55
3.6 Some Allophony in other languages . . . . . . . 58
3.7 Universal Properties and Language-Specific Properties . . . . . . . . . . . . . . . . . . . . . . . . . 59
3.8 Summary . . . . . . . . . . . . . . . . . . . . . . . . 60
3.9 Hints on solving phonology problems . . . . . . 61
3.10 Exercises . . . . . . . . . . . . . . . . . . . . . . . . 64


## 3.1 Further Phonetic Details

The symbols of the IPA we have learned allow us to represent in much greater detail the speech of most people with English-type grammars than would ever be possible using the English writing system alone, which is unsystematic, arbitrary, and confusing. The IPA also includes symbols to transcribe the basic segments of all other known human languages. However, as linguists studied in greater detail the sounds of even well-known grammars, such as those of the English type, they discovered that there were small details of articulation which the IPA alone failed to capture. For example, depending upon the dialect, speakers pronounce [æ] differently, sometimes further front and sometimes further back. The IPA has been supplemented with many additional symbols (which, fortunately, you don't have to learn, even though we did). So [æ̟] indicates, through the use of a **diacritic** (in this case, a subscripted plus-sign), a further front low front unround vowel such as is found in some pronunciations of words such as 'cat.' The contrast between a 'more rounded' and a 'less rounded' type of [o] can be indicated as well ([o̜] for the less rounded and [o̹] for the more rounded). A voiceless velar stop produced with lip-rounding, as in the normal pronunciation of a word like 'cool,' is indicated as follows: [k̫ul]. The set of possible symbols increased dramatically as phoneticians made more and more precise observations about the details of articulation.

With the advent of the use of computers to analyze the acoustic properties of human speech, however, it quickly became apparent that the goal of having a different symbol for each 'speech sound' could never be attained. The computer revealed that, in fact, each of the [æ]'s that a single speaker produces (if one were to record them saying, e.g., [kæt] ten times) is slightly different. We would need an infinite number of symbols with diacritics to represent all the details of speech, and that's a *lot* of symbols...

It became clear that a better goal was not to pay attention to *all* of the differences between the [æ]'s or [k]'s that speakers produced, but to focus on differences that might be used in *some* human language to indicate a change in meaning. An example might help with this. Speakers with English-type grammars normally produce round consonants before round vowels (remember, *all* back vowels are also round in English-type grammars). So we say [k̫ul] 'cool,' [t̫u] 'two,' and [m̫un] 'moon.' However, if you were to say [kul] or [mun] (with no rounding of the consonant), everyone would know that you meant 'cool' and 'moon,' respectively. In the case of speakers with Marshallese-type grammars, however, the difference between round and unround consonants is very important. For example, [**am nik̫nik̫**] means 'your cloth' but [**am̫ nik̫nik̫**] means 'our (i.e., belonging to me and some other people, but *not* you) cloth.' Marshallese speakers thus have to pay attention to the rounding of [m] (and of other consonants as well), whereas speakers of English do not. Since Marshallese [am] 'your' and [am̫] 'our (excluding you)' differ *only* in the lip-rounding that we hear on the labial nasal, any speaker of Marshallese who did not attend to this contrast would be very confused (and very confusing).

On the other hand, the difference between [ʃ] and [ʧ] is of no consequence in Marshallese. For example, the verb which means (according to the dictionary) ‘to go barefoot’[1] can be pronounced [ʃintəb̫] or [ʧintəb̫] – no Marshallese speaker will pay any more attention to this difference than you would to the difference between [k̫ul] and [kul].

The contrast between round and unround [m] or [k] is thus very different from the distinction between a slightly fronter [æ] (i.e., an [æ̟]) and a slightly backer one (indicated in the IPA by [æ̠]). *No* human language uses the difference between these [æ]’s to signal a difference in meaning, in the way Marshallese uses the [m]:[m̫] contrast, or English the [ʃ]:[ʧ] one.

The second important implication of the difference between the status of rounding or affricates in English and Marshallese is that it appears that each grammar uses some set of features of the speech stream to pay attention to and ignores others (though the ones they ignore may be just as present as those being attended to, as is the case with consonant rounding in English or the [ʃ]:[ʧ] variation in Marshallese). It appears, then, that part of what you *know*, as a speaker with a particular grammar, is which aspects of the speech stream matter for conveying differences in meaning and which ones do not.[2] Since the knowledge of what to ignore and what to pay attention to appears to vary from grammar to grammar (so that, e.g., Marshallese-type grammars differ from English-type grammars in this regard), you must *learn* to ignore certain aspects of the speech signal when you learn your grammar. The next section will explore the nature of the knowledge you acquire with respect to this.

## 3.2 The Phoneme

Imagine that you are about to get married (or, if that makes you too nervous, or seems too unlikely, imagine that you are about to attend someone else’s wedding – if that too is beyond the scope of your imagination, you will have to develop your own analogy). The wedding ceremony – to be a real wedding – must have certain features. For example, someone other than the

[1] This verb also means ‘to eat only one food.’

[2] Like all linguistic knowledge, this knowledge is, unfortunately, *tacit*, or hidden from consciousness.

two people getting married must be present (to perform the ritual ceremony). Depending on your culture, various other things will have to happen if the two people are to be truly married. One could say that the things that determine whether or not what happens is actually a wedding (like is there someone there for you to marry, is there someone there to perform the ceremony, is that person lawfully empowered to marry a given couple, etc.) are *criterial* as to whether or not a wedding has taken place. In addition to these criterial features, every wedding will have many many features which are *incidental* – i.e., irrelevant to the question of whether or not what took place was actually a wedding. For example, the weight of the groom does not determine whether or not a wedding has taken place, nor does the time of day or year, the size of the bride's feet, the number of people named 'Bob' who attend, or an infinite number of other aspects. Note that every wedding has these incidental features – at every wedding the groom weighs some amount, it takes place at a given time of day and year, the bride's feet have a certain dimension, there are some number ($\geq 0$) of people named 'Bob,' etc.

Similarly, we could say that in English-type grammars the difference between [ʃ] and [ʧ] is *criterial* for the determination of meaning, whereas the difference between [m] and [m̫] is *incidental* – i.e., present but not decisive for whether or not the word 'moon' is interpreted by a speaker of English to be the round white object which sometimes appears in the sky at night. By contrast, in Marshallese the difference between [ʃ] and [ʧ] is *incidental* to the determination of the meaning of a given string, whereas the difference between [m] and [m̫] is *criterial.* Speakers must produce and pay attention to those segments which determine meaning differences (otherwise they won't know what anyone is saying) but can safely ignore those differences which do not influence meaning (who cares if you say the [m] of 'moon' with lip-rounding or not? it's going to mean 'moon' either way). The sound differences that matter for a given grammar are called **phonemic** differences. Differences that do not matter for the purpose of determining meaning (but are present, like the dimensions of the bride's feet at the wedding) are called **phonetic** differences. The set of sounds of a given grammar, defined in such a way as to list only those aspects of their production that are *criterial* for the determination of meaning, are known as the **phonemes** of that language. Thus an /m/[3] which has no indication as to whether or not it

[3]Phonemes are placed between slashes, to keep them clearly distinct from phonetic segments, which are in square brackets.

involves lip-rounding is a phoneme of English, whereas in Marshallese /m/ (an *unrounded* labial nasal) and /m̫/ (a *rounded* labial nasal) are separate phonemes. Phonemes are what one must learn to pay attention to in order to ultimately derive accurate semantic representations.

## 3.3 Minimal Pairs & Allophones

How can we determine what the set of phonemes for a given grammar are? One technique which works quite well is that of *substitution*. The idea is quite simple. Since, in English [bæt] means 'bat' and [pæt] means 'pat,' the difference between [b] and [p] is a **phonemic** one. This is shown by the fact that if we substitute the [b] of [bæt] with [p], we get a difference in meaning. Words which differ in meaning and have that difference signalled by just a single segment are called *minimal pairs* (meaning 'minimally-distinct pairs' of words). Since the only difference between [bæt] and [pæt] resides in the [b]:[p] contrast, we can infer that these two segments are used to signal meaning differences. By contrast, if we substitute an unrounded [m] for the [m̫] of [m̫un] 'moon,' giving [mun], the newly created word may sound a little funny, but it still means 'moon.' The difference between [m] and [m̫] is therefore *not* a phonemic one.

Distinctions which are merely phonetic (rather than phonemic), although they do not play any role in determining meaning, are generally found to have a well-defined distribution. For example, if you isolate the vowel of 'men' (by starting out to say 'men' but then, instead of producing the final nasal, simply prolonging the vowel) you will discover that this vowel is, for speakers with English-type grammars, *nasalized* (i.e., the velum is lowered as it is for nasal consonants). It contrasts, therefore, with the vowel of a word such as [mɛt] 'met' which (as you will discover if you isolate it using a similar technique) is not nasalized. That is, if we want to be phonetically precise, 'men' should be transliterated [mɛ̃n]. (The 'tilde' diacritic indicates nasalization of vowels.) In the output of English speakers, we find both [ɛ] (in 'met,' 'bet,' etc.) and [ɛ̃] (in 'men,' 'Ben,' etc.). Similarly, in the speech of speakers with French-type grammars we find both these vowels, [ɛ] in words like [fɛ] 'fait' (fact), [ɛ̃] in words like [fɛ̃] 'fin' (end). Notice that the same segments are contained in the phonetic output of both types of speakers. However, as the examples given above show, in French if you substitute [ɛ̃] for the [ɛ] of 'fait', you get [fɛ̃] 'fin,' which means something completely different. However, if you substitute the [ɛ̃] of 'men' into 'met' in English, the resulting [mɛ̃t] may sound a little odd but the listener will still understand you to have said the past tense of the verb 'meet'. They may infer that you have a cold or that you are performing that phonetic experiment which involves placing an olive up your nose but no one would think you said anything but 'met' (i.e., the substitution does *not* lead to a change in meaning, unlike in French). While French [fɛ] and [fɛ̃] form a **minimal pair**, we cannot construct two words with different meaning in English using the [ɛ]:[ɛ̃] contrast.

In fact, all English vowels are nasalized before nasals and *only* before nasals (barring some physical – and therefore non-linguistic – problem, such as a cold). In French, as [fɛ̃] 'fin' shows, this same statement is not true (we find nasalized vowels which are not before nasals). Since the difference between [ɛ] and [ɛ̃] is a phonemic one in French-type grammars, we refer to [ɛ] and [ɛ̃] as *phonemes* of French.

But what is the relationship between nasalized and non-nasalized vowels in English-type grammars? As we've seen, since the distinction cannot express a difference in meaning, there is no question of the nasalized and non-nasalized vowels being separate phonemes. The distribution of nasalized vowels is, in English but not in French, completely predictable. We can state the prediction in the form of the following 'rule' of grammars of the English-type:

(1) Rule 1: Nasalize vowels before nasals.

This rule[4] accounts for the fact that in English nasalized vowels occur always and only before nasals. The vowel of [mɛt] will not get nasalized by Rule 1, for example, since it is not before a nasal. On the other hand, imagine that in your brain you store the word 'men' as /mɛn/ (without nasalization). When you run /mɛn/ through the grammar, a process such as that in (2) will take place.

(2) /mɛn/ $\Longrightarrow$ (via Rule 1) [mɛ̃n]

The stored form /mɛn/ is converted by the application of (1) into a form

[4]Note that we are using the term 'rule' in the standard way it is used in linguistics. It is quite different from a rule like 'No pushing at the drinking fountain' which one can follow or not follow as one chooses. Instead it represents the *knowledge* that a speaker has which gives rise to the distribution of phonetic features such as vowel nasalization.

with a nasalized vowel, which you then send on to the articulatory system and ultimately pronounce with nasalization.

As we pointed out above, English-type grammars also generally round consonants which occur immediately before round vowels. We can state this 'consonant rounding' rule as in (3).

(3) Rule 2: Make consonants round before round vowels.

Now imagine that we have 'moon' stored as /mun/. To derive the actual pronounced form of this word, the grammar applies its **phonological rules** to the phonemic form /mun/ which is the input to the grammar. The result is a **phonological derivation**, such as that in (4) for 'moon.'

(4) /mun/ $\Longrightarrow$ (via Rule 1) [mũn] $\Longrightarrow$ (via Rule 2) [m̫ũn]

The final form has both nasalization of its vowel (by Rule 1) and lip-rounding of its consonant (by Rule 2). Note that since Rules (1) and (2) apply to all input forms to the grammar, the output of these rules, taken together, will create a situation in which [ɛ̃] will occur only where [ɛ] does not (before nasals), and vice-versa. The two phonetic segments [ɛ̃] and [ɛ] are in **complementary distribution** – one never occurs in the same **environment** as the other.

This type of analysis carefully distinguishes between what kind of information you store in your head about the sounds in a word and what kind of instructions you send to your articulators in order to get them to produce the word. At the level of mental storage, you need only store enough information about 'men' to distinguish it from 'mean,' 'moan,' 'moon,' 'man,' 'met,' 'mess,' etc. Since nasalization of vowels and rounding of consonants is never relevant to keeping meanings distinct in English-type grammars, there is no reason to store nasalization and consonant rounding information in our brains about English words. Nasalization and consonant rounding are *predictable* aspects of English-type grammar output. This contrasts sharply with the cases of French (with respect to nasalization) and Marshallese (with respect to consonant rounding). If we did not store nasalization information in French, [fɛ̃] and [fɛ] would both be stored /fɛ/. Similarly, if we did not store consonant roundedness information in Marshallese, [am̫] 'our' and [am] 'your' would both be stored as /am/. Clearly we cannot write straightforward rules for the stored forms of these French and Marshallese words which would insert nasalization and rounding in just the right words.[5]

The segments used in our mental dictionary are the phonemes of the language. These phonemes can be realized, as a result of the application of phonological rules such as those in (1) and (2), in a variety of ways. So, for example, the phoneme /m/ in English can be realized as [m̫] (before round vowels) or as [m] (elsewhere), given Rule 2. The phoneme [ɛ] can be realized as [ɛ̃] (before nasals) or as [ɛ] (elsewhere), given that English grammars have Rule (1). The various phonetic realizations of a phoneme (and we have not given anything like a complete list here) are called the **allophones** of that phoneme. So, in English-type grammars, the phoneme /ɛ/ has the allophones [ɛ̃] and [ɛ] (at least – as we noted, the list is not complete), while the phoneme /m/ has at least the allophones [m̫] and [m]. It will be useful to think of the phonological component as consisting of two levels of abstract mental representations with a computational system that 'converts' one type of representation into another. The most abstract level is the level of the phoneme, or the **underlying representation** (UR) and the less abstract is the phonetic level, or **surface representation**. In the production of speech, the phonological computation system converts underlying (phonemic) representations into surface (phonetic) representations. These surface representations are then further converted into articulatory gestures (a process *external* to the grammar). In the comprehension of language, relevant auditory input (speech sounds) is converted into surface (phonetic) representations (again, a process *external* to the grammar). The surface representation is then converted by the phonological component into an underlying (phonemic) representation. Again, *unpredictable*, and therefore important to remember, information is stored only at the underlying or phonemic level.

## 3.4 More English allophony

When a native speaker with an English-type grammar says the word 'pick' one can observe a burst of air following the release of the stop closure of the initial segment. You can test this out by putting a candle, match, or lighter flame in front of your lips when you say 'pick' (not too close). The flame

[5]Obviously, we can posit a rule which, for French, says 'nasalize the [ɛ] in the word for 'end,' but notice that since the rule explicitly mentions the word that it applies to, we might just as well posit the nasalization as an inherent property of the word and get rid of the 'rule' altogether.

will flicker dramatically after the release of the stop. This burst of air is known as 'aspiration.' All English voiceless stops are aspirated when they are the initial segment of a stressed syllable. Thus 'pick', 'tick', and 'kick' are pronounced, by native English-type speakers, as [pʰɪk], [tʰɪk], and [kʰɪk].[6]

The aspiration of voiceless stops is *not* found when the voiceless stop is not syllable-initial. So, e.g., in words like 'spat', 'stat', and 'scat' there is no aspiration ([spæt], [stæt], and [skæt]). The aspiration of voiceless stops is thus predictable – a voiceless stop will be aspirated if it is the first segment of a stressed syllable, otherwise it will not. English-type grammars thus have a rule such as in (5) below.

(5) Rule 3: Aspirate voiceless stops at the beginning of stressed syllables.

We can thus say that the phonemes /p/, /t/, and /k/ have allophones as in (6-8).

(6) /p/ has the allophone [pʰ] at the beginning of stressed syllable, the allophone [p] elsewhere

(7) /t/ has the allophone [tʰ] at the beginning of stressed syllable, the allophone [t] elsewhere

(8) /k/ has the allophone [kʰ] at the beginning of stressed syllable, the allophone [k] elsewhere

It should be clear at this point that allophones are *generated* by phonological rules.

One more case of allophonic alternation in English-type grammars should suffice to make the principle of the phoneme clear to you. If you say the words 'bat' and 'bad' to yourself, you will notice that there is in fact a significant difference in the length of the two vowels. The vowel of 'bad' is considerably longer than the vowel of 'bat.' It turns out that in English-type grammars, vowels are much longer before voiced consonants than they are before voiceless consonants. Vowel length in the IPA is indicated by a colon-like symbol [ː]. Thus, 'bad' would be transliterated [bæːd], while 'bat' is [bæt] (no length-mark).[7] English-type grammars thus have a rule such as

[6] A superscripted 'h' is the diacritic for aspiration.

[7] This notion of 'length' should not be confused with the somewhat ideosyncratic use of this term in traditional English grammar, where the diphthong [aj] is called 'long i' and other such peculiarities. A long vowel is, not surprisingly, a vowel that lasts longer than a

that in (9).

(9) Rule 4: Lengthen vowels before voiced consonants.

The application of this rule to a stored lexical item such as /bæd/ (for which no length is stored, since length is predictable) produces the output representation [bæːd].

We can see the combined effects of the phonological rules we have posited for English-type grammars if we examine the **phonological derivation** of a word such as /konhɛd/ 'conehead.' The stored lexical item – with all of the predictable information stripped out – is /konhɛd/, as the phonemic bracketing indicates. Rule 1 says that vowels before nasals must be nasalized in English-type grammars. It therefore converts /konhɛd/ to [kõnhɛd]. Rule 2 says that consonants are rounded before round vowels, of which [õ] is one. Rule 2 therefore turns [kõnhɛd] into [k̫õnhɛd]. Rule 3 says that voiceless stops at the beginning of stressed syllables are aspirated. It therefore converts this form into [k̫ʰõnhɛd]. Finally, according to Rule 4, vowels are lengthened in English-type grammars before voiced consonants. Rule 4 thus changes our form into the actual output form: [k̫ʰõːnhɛːd].

## 3.5 Rule ordering

So far, the rules we have examined have not required any explicit ordering. For example, it matters little to the final outcome whether we apply Rule 1, which nasalizes vowels, before or after we apply Rule 4, which lengthens vowels before voiced consonants. If we apply Rule 1 first, we'll get a nasalized vowel, which, being a vowel, will lengthen by Rule 4. If we apply Rule 4 first, on the other hand, we'll get a lengthened vowel. Rule 1 will nasalize this lengthened vowel (since it nasalizes *all* vowels), and the result will be exactly the same.

In some cases, however, explicit ordering is required in order to derive the correct result. Examine the alternations in (10ab).

(10) a. 'fat' [fæt] : 'fattest' [fæɾəst], 'hit' [hɪt] : 'hitting' [hɪɾɪŋ]
b. 'wide' [wajd] : 'widest' [wajɾəst], 'kid' [kɪd] : 'kidding' [kɪɾɪŋ]

Since in their simple form 'fat' ends in a /t/ and 'wide' in a /d/, it seems

---

short vowel but has *no difference in quality.*

clear that we need a rule which converts both /t/ and /d/ to [ɾ] under certain conditions. The details need not detain us too seriously at this point, so we will state the rule as in (11).

(11) Turn an alveolar stop into a voiced, alveolar flap ([ɾ]) when it is found between two vowels and at the beginning of an unstressed syllable.[8]

In /fæt/ the /t/ is not at the start of a syllable at all (it is at the end of a stressed syllable), as is the /d/ of /wajd/. For this reason, these segments surface as [t] and [d], respectively. However, in 'fattest' and 'widest' the /t/ and /d/ have become the initial segment of the second syllable, which is unstressed.[9] They therefore undergo the 'flap' rule, taking the form indicated in (10ab). This rule is pervasive in most varieties of North American English (though it is rare in grammars of the English type found in England).

Now examine the date in (12ab).

(12) a. [ajz] 'eyes', [lajz] 'lies', [trajd] 'tried', [trajb] 'tribe'
b. [əjs] 'ice', [ləjs] 'lice', [trəjt] 'trite', [trəjp] 'tripe'

The diphthong found in the (b) examples in (12) is clearly different from that found in the (a) examples. In fact, many North American varieties of English, particularly those found in and near to Canada, have the rule in (13), which triggers the different realizations of /aj/.

(13) /aj/ becomes [əj] before voiceless consonants

This rule is responsible for the data in (12) as well as for alternations such as [flaj] 'fly' vs. [fləjt] 'flight.'

Given the 'aj-raising' rule of (13), 'write' will have the underlying form /rajt/ and 'ride' will have the underlying form /rajd/. The fact that /rajt/ will be output as [rəjt] is *predictable* (since /t/ is voiceless) and therefore will not be stored in the mental lexicon. What happens when we add the agent suffix /r̩/ to these verbs? We get phonemic representations that have the form /rajtr̩/ 'writer' and /rajdr̩/ 'rider' (notice that these forms have two syllables – with stress on the first syllable – whereas the verbs were

[8] *Both* of these conditions must be true and the term 'vowels' here includes all segments that 'act like' vowels for purposes of syllable structure, thus syllabic liquids and nasals count as vowels.

[9] Rather than go through a lengthy explanation of syllable structure here, we ask that you trust us on the relative positions of segments within syllables.

monosyllabic). The conditions for both the Flapping and the Raising rules are met by the form /rajtr̩/ – the /t/ is at the start of the second, unstressed syllable, and the /aj/ is before a voiceless consonant. What order do we apply the rules in? (14a) shows the result of applying the Flapping rule before the Raising rule, (14b) shows the results of applying the rules in the opposite order.

(14) a. /rajtr̩/ > (Rule 11) [rajɾr̩] > (Rule 13) [rajɾr̩] (no change, because [ɾ] is *voiced* and 'aj-raising' applies only before voiceless consonants)
b. /rajtr̩/ > (Rule 13) [rəjtr̩] > (Rule 11) [rəjɾr̩]

Note that the output of (14a) is identical to the output we get for 'rider': /rajdr̩/ > (Flapping rule) [rajɾr̩]. However, 'rider' and 'writer' are not homophonous (i.e., they don't sound alike). 'Writer' has the [əj] diphthong in it, even though the flap which follows the diphthong is *voiced*! Therefore, (14b) is the correct result. We must specify that the Raising rule applies *before* the Flapping rule. As this example shows, certain phonological rules which generate allophones must be ordered with respect to one another (Flapping and Raising), whereas the ordering of certain other rules with respect to one another is not significant (as we saw with nasalization and lengthening of vowels). When you are solving a phonology problem, you must always ask yourself whether or not the rules you have posited require ordering.

You may have noticed that the phonological rules which we have posited apply to *classes* of segments. Rule 1 applies to *vowels* before *nasals*, Rule 2 involves *consonants* and *round vowels*, Rule 3 applies to all *voiceless stops*, Rule 4 before *voiced consonants*, etc. We do not generally find rules applying before random sets of sounds – sounds which share nothing in common with one another. So, for example,a rule like 'lengthen vowels before /b/, /s/, and /l/' would be quite odd. The classes to which rules apply are generally known as **natural classes** of segments – you should always try to state any observed phonological processes in terms of the properties of the set of segments involved rather than by simply listing the segments. Thus 'vowels are nasalized before nasals' is a better scientific hypothesis than 'vowels are nasalized before /m/, /n/, and /ŋ/', even though in English-type grammars the two statements produce the same results (since /m/, /n/, and /ŋ/ are the only nasals in English). 'Better scientific hypotheses' have the following properties: 1) they are the most general statements possible which make no false predictions; 2) they are stated in the most economical fashion; 3) they are based on empirical evidence; and 4) they are testable (and thus falsifiable).

Many phonological rules have the result of making one segment more like an adjacent segment. This is true of both the nasalization and rounding cases we have just discussed. This process is called **assimilation**. There are, of course other types of rules – rules which delete segments, rules which insert segments, rules which make a segment *less* like an adjacent segment (dissimilation), and so on. Assimilation rules are by far the most common, however.

## 3.6 Some Allophony in other languages

The principles developed on the basis of English data above are, of course, general linguistic principles which can be used in the analysis of any language. All languages have phonological rules. The specifics of these rules differ from language to language but the general structure of the rules is the same. Because the basic structure of phonological rules is the same in all natural language, the same principles can be used to determine the specifics of phonological rules in particular languages. For example, examine the data in (15) from Malay (spoken in Malaysia and Singapore).[10]

(15) a. [kərəta] 'car'
b. [mãkan] 'eat'
c. [mãrah] 'scold'
d. [mə̃laraŋ] 'forbid'
e. [rumãh] 'house'

Is nasalization predictable in Malay? [Take a moment to think about it – you might want to look at the data to help you decide. . . ] Yes! Malay appears to have a rule like that in (16).

(16) Nasalize all vowels which immediately follow a nasal.

The *phonemic* representation of these words in the minds of Malay speakers are thus those in (17). (Recall that phonemic representations have redundant,

[10] As is normal in such examples, the data is presented only to reveal a principle of analysis. The actual facts of nasalization in Malay are slightly more complicated.

predictable information stripped out of them.)

(17) a. /kərəta/ ‘car’
b. /makan/ ‘eat’
c. /marah/ ‘scold’
d. /məlaraŋ/ ‘forbid’
e. /rumah/ ‘house’

If we apply the rule in (16) to the forms in (17) we will get the data in (15), as desired.

Now examine the Khmer phonetic data in (18).

(18) a. [pɔːŋ] ‘to wish’
b. [tɔp] ‘to support’
c. [kat] ‘to cut’
d. [pʰɔːŋ] ‘also’
e. [tʰɔp] ‘to be suffocated’
f. [kʰat] ‘to polish’

Is aspiration predictable in Khmer, as it is in English? [We’ll put this here to give you time to think about it before you answer. . . ] No! /p/ and /pʰ/ are *separate phonemes* in Khmer (even though in English they are allophones of /p/). If they were allophones of /p/ as in English, the words ‘to wish’ and ‘also’ would both be stored in the lexicon as /pɔːŋ/. But notice there is *no way* to write a rule (without mentioning the specific forms involved, which makes the rule useless) which would cause this form to surface with aspiration in ‘also’ but without aspiration in ‘to wish.’ The phonemic forms of these words in Khmer must therefore contain specifications as to whether or not a given voiceless stop is aspirated, while in English no such specification is required (because aspiration is predictable).

## 3.7 Universal Properties and Language-Specific Properties

If the human linguistic system is innate, we would expect many properties of the system to be **universal**. Indeed, all human languages have a phonological component where abstract representations of sound are processed. On the other hand, it is fairly clear that some part of human linguistic knowledge is acquired through exposure to a particular language environment and is thus language-specific. One of the central questions in linguistics, then, is 'What is a *possible* human linguistic system?' Note that the answer to this question will necessarily include both universal *and* language-specific properties. So while all languages have phonological rules, not all languages have a phonological rule which causes voiceless stops to be aspirated in certain positions. The specifics of rules can only be determined through examining the relevant data.[11] There is another aspect of phonology which falls into the language-specific realm. In Chapter 1, we noted that there were certain combinations of sounds, certain 'words,' that native speakers would immediately identify as 'not possible words.' For English, we gave examples such as 'ngloopi' and 'dnkli.' All native speakers have intuitions of this type – what sounds can be put together, what combinations of sounds can form a syllable, etc. – about their native language. These language-specific properties are called the **phonotactics** of the language. So 'ngloopi' violates a **phonotactic constraint** of English – English does not allow the sequence 'ngl' at the beginning of the syllable. Recall, however, that simply because it does not occur in English, does not mean that it cannot occur in some other language. It *may* be the case that this is a universal restriction on language (i.e., not a possible syllable-initial sequence in human language) but the only way to determine whether it is actually the case is through examination of the relevant empirical evidence. (It turns out that other languages *do* have initial clusters of that type.)

## 3.8 Summary

We have seen in this section that the lexicon (the mental dictionary) stores only *unpredictable* information about the sounds in particular lexical items in a language. This requires the notion of the *phoneme* or *underlying representation* – which is the name for the segments used in our mental dictionary – and the notion *allophones* or *surface representations* – which is the name for the various phonetic realizations of those phonemes. The various realizations themselves are generated by the application of *phonological rules* – sometimes ordered, sometimes not – to the representations stored in the mental lexicon. Grammars differ both with respect to their phonemes and their phonolog-

[11] This means that you should *not* carry over what you know about the specifics of English phonological rules to other languages.

ical rules. Because their phonological rules differ, different grammars may produce different allophones for a given underlying segment.

## 3.9 Hints on solving phonology problems

The basic question is always the same: What sorts of things can a human know with respect to the sound system of natural language? We approach this question by looking at *phonetic* data from a 'language' and then we try to figure out the unconscious system that produces such forms.

The various kinds of detailed questions that we ask are the following. Are two or more segments in complementary distribution (i.e. never occur in the same environment) and are they thus allophones of a single phoneme? Or, are the segments in question separate phonemes (listed distinctly in the phonological inventory - i.e. the underlying/'mental' representation of sounds in a particular language)? Note that if you are told to consider more than two segments, it may be the case that only two out of three are allophones (or, of course, that none are).

A somewhat mechanical approach to answering this question involves the following steps. Step (1): List all the environments. (The environments are just the 'surroundings.') That is, mechanically list all segments (including word boundaries, as well as tone or syllable stress if they look like they might be important) which occur to the immediate left and to the immediate right of the segments in question. Step (2): Consider the left side. If there is no overlap, e.g. the environment on the left is never the same for Segment A as it is for Segment B, then they are in complementary distribution and you have solved the problem. If there *is* overlap, then follow the same step for the right side. (You should keep in mind during this process that when two items are allophones, they will share a large number of phonetic features in common although 'trick' problem sets, like whether /h/ and /ŋ/ are allophones in English rarely occur.) If you have found nothing interesting in the way of complementary distribution on either the left or right sides, proceed to Step (3). Step (3) is to consider that the environment could include *both* sides – that is, consider the fact that a *combination* of left and right adjacent segments is the environment. A hypothetical example would be that Segment A is only found between two /i's/. Alternately, you may need to view things from a slightly broader perspective such as whether Segment A occurs only between two vowels. Or, you may need to consider things such as tone or syllable stress. After this examination, if you find that there is no overlap in environment, you may conclude that the segments are allophones. Otherwise, (if you consistently find an overlap) they are phonemes.

Writing up the solution involves formalizing your findings, to some extent. If you conclude that two or more segments are allophones of a single phoneme, you need to decide which of the allophones you will choose to represent the phoneme, or underlying representation (UR). Since we assume that language is represented economically in the minds of speakers, we choose the segment which is found in the *greatest variety* of environments as the UR/phoneme. This is the 'elsewhere' segment. As an illustration, we will use nasalization of vowels in English which we discussed earlier. In this case, the non-nasal (oral) vowels occur in the *greatest variety* of environments – everywhere except before nasals.[12] Non-nasal vowels are thus the 'elsewhere' case and should be chosen as the phonemes. If we were to choose nasal vowels as the phonemes, we would have to list as the environment for the rule every segment in the language apart from the three nasal consonants. This would be both uneconomical and also miss a generalization about phonological processes since nasalization of vowels before nasals is clearly an assimilation process.

Linguists have developed a 'shorthand' for writing rules which is in widespread use. A basic understanding of the form of these statements will be a useful thing for you to acquire. Basically, phonological rules have the form in (19).

(19) A → B / X__

This formalism is to be read 'A becomes (→) B in the environment (/) in which A follows X. The '__' tells you where the affected segment is in the string if the rule is to apply to it. Thus, the nasalization rule would be rewritten as in (20).

(20) vowel → nasal vowel / __ nasal

This means vowels are nasalized before nasals. The Malay vowel nasalization rule in (16), by contrast, would be written as in (21).

(21) vowel → nasal vowel / nasal __

[12]This is true because oral vowels occur before stops, affricates, fricatives, word boundaries, voiced sounds, voiceless sounds, and so on. Nasal vowels, on the other hand, only occur before nasal consonants, a much more restricted set than the oral vowel set and one which, in addition, forms a natural class.

This means that vowels are nasalized *after* nasals. If you had a language which nasalized vowels only when the vowel was *between* two nasals, the rule would have the form in (22).

(22) vowel → nasal vowel / nasal __ nasal

This rule specifies both a preceding and a following nasal being required to trigger nasalization of vowels.

Note that for both the English and Malay cases, we do not list every vowel and its nasal counterpart individually in the statement of the rule. Instead, we have made the most general statement possible which still captures the data. If a rule applies to all vowels, then there is no reason to be more specific than just stating 'vowels.' The vast majority of rules apply to natural classes of segments and natural classes may be easily described using only a few features, e.g., high vowels; labials; velars; stops; lax vowels; voiced consonants; and so forth.

If you are asked to do a derivation as part of your solution:

A derivation consists of:

- the underlying form (this should always include the segment that is in the speaker's phonological inventory, i.e. the phoneme or UR and should appear in *slashes*);
- the rule/rules (Note: if there are more than one rule, they may have to be ordered with respect to one another – recall the English Flapping and Raising rules.)
- the surface form (in square, phonetic brackets).

If the word [bĩn] is given in the data for nasalized vowels in English, a derivation of [bĩn] would look like this:

$$\begin{array}{c}\text{/bin/}\\ \Big\downarrow \text{ nasalization rule}\\ \text{[b}\tilde{\text{i}}\text{n]}\end{array}$$

## 3.10 Exercises

1. Mokilese (an Oceanic language spoken on Mokil). Examine the following data, noting the distribution of voiceless (with a ring under the vowel) and voiced vowels.

1. [pi̥san] ‘full of leaves’
2. [uduk] ‘flesh’
3. [kaskas] ‘to throw’
4. [poki] ‘to strike something’
5. [tu̥pu̥kta] ‘bought’
6. [pu̥ko] ‘basket’
7. [ki̥sa] ‘we two’
8. [pil] ‘water’
9. [apid] ‘outrigger support’
10. [luʤuk] ‘to tackle’
11. [su̥p̫o] ‘firewood’
12. [kam̫ɔki̥ti] ‘to move’

a. Which rule accounts for the voiceless vowel allophones in the data above?

(a) Vowels become voiceless before voiceless stops.
(b) High vowels become voiceless between voiceless consonants.
(c) High vowels become voiceless after voiceless consonants.
(d) Vowels become voiceless between voiceless consonants.

b. What should the phonetic realization of Mokilese /pisa/ ‘fish body part (from the start of the tail to just below the gills)’ be?

(a) [pisḁ]

(b) [pisa]

(c) [pi̥sa]

(d) [pi̥sḁ]

c. The phonemic representation of Mokilese ‘bought’ is:

(a) /tu̥pu̥ktḁ/

(b) /tupuktḁ/

(c) /tu̥pu̥kta/

(d) /tupukta/

2. Canadian French. Examine the data below. [i] and [ɪ] are allophones of one phoneme, as are [y] (a tense high front rounded vowel) and [ʏ] (a lax high front rounded vowel), [e] and [ɛ], [o] and [ɔ], and [u] and [ʊ].

1. [pilʏl] ‘pill’
2. [fyme] ‘smoke’
3. [griʃe] ‘to crunch’
4. [lynɛt] ‘glasses’
5. [grɪʃ] ‘it crunches’
6. [frole] ‘to skim’
7. [pətsi] ‘little (masc.)’
8. [pətsɪt] ‘little (fem.)’

9. [pɔrt] 'door'
10. [vitamɪn] 'vitamin'
11. [saly] 'hi'
12. [ʒʏp] 'skirt'
13. [bote] 'beauty'
14. [bɔt] 'boot'
15. [fo] 'false'
16. [tɔrdzy] 'twisted'
17. [zero] 'zero'
18. [pɔm] 'apple'
19. [lʏn] 'moon'
20. [plʏs] 'more'
21. [pɪp] 'pipe'
22. [grimas] 'grimace'
23. [tʊʃ] 'touch'
24. [ru] 'wheel'
25. [suvã] 'often'
26. [trupo] 'herd'
27. [sʊp] 'flexible'
28. [tryke] 'to fake'
29. [fini] 'finished'
30. [rʊt] 'road'

a. What is the rule that governs when the tense ([i], [y], [o], [u]) and lax ([ɪ], [ʏ], [ɔ], [ʊ]) allophones appear?

(a) The lax variants are found between consonants, the tense variants elsewhere.

(b) The lax variants are found before word-final consonants, the tense variants elsewhere.

(c) The lax variants are found in closed syllables (i.e., when a consonant follows them which is not itself followed by a vowel), the tense variants elsewhere.

(d) The lax variants are found before voiced consonants, the tense variants elsewhere.

b. The phonemic representation of the word for 'glasses' (4) in this language is:

(a) /lʏnɛt/

(b) /lynɛt/

(c) /lynet/

(d) /lʏnet/

3. Tamil (Dravidian, South India). Examine the data below from the Tamil language, paying particular attention to the initial glides ([w] and [j]). [Note that [ɖ] is a voiced, retroflex stop, produced with the tip of the tongue pointed towards the back of the mouth – this fact is not relevant to the solution for this problem.]

1. [jeli] 'rat'
2. [woɖi] 'break'
3. [arivu] 'knowledge'

4. [jiː] ‘fly’
5. [woːlaj] ‘palm leaf’
6. [ain̪t̪u] ‘five’
7. [jilaj] ‘leaf’
8. [wuːsi] ‘needle’
9. [aːsaj] ‘desire’
10. [jeŋgeː] ‘where’
11. [wujir] ‘life’
12. [aːru] ‘river’
13. [jiɖuppu] ‘waist’
14. [woːram] ‘edge’
15. [aːdi] ‘origin’

a. Which of the following best accounts for the data above?

(a) The initial glides are not predictable.

(b) The initial glides are predictable. The initial round glide [w] appears before non-round high and mid vowels; the initial non-round glide [j] appears before round high and mid vowels.

(c) The initial glides are predictable. The initial round glide [w] appears before round high and mid vowels; the initial non-round glide [j] appears before non-round high and mid vowels.

(d) The initial glides are predictable. The initial round glide [w] appears word-initially before high vowels; the initial non-round glide [j] appears before non-high vowels.

b. What are the phonemic representations of ‘rat’ and ‘life’?

(a) /eli/ and /ujir/

(b) /weli/ and /wujir/

(c) /jeli/ and /wujir/

(d) /jeli/ and /jujir/

4. West Tarangan (Maluku, Indonesia). Examine the data below from the West Tarangan language. The voiceless, bilabial fricative /ɸ/ (like an [f], but without the involvement of the teeth) has two allophones: [ɸ] and [p].

1. [ɸit] 'night'
2. [ɸɔp] 'pig'
3. [ɸaɸə] 'ground'
4. [ɛtaɸakə] 'it capsizes'
5. [ɔlɸɔt] 'egret'
6. [gurep] 'nail'
7. [nakɸɔ] 'I will carry'
8. [tɔp] 'short'

a. Which of the following statements accounts for the distribution of these allophones?

(a) You can't tell when [ɸ] will appear and when [p] will appear.

(b) /ɸ/ → [p] at the ends of words, [ɸ] elsewhere.

(c) /ɸ/ → [p] after vowels, [ɸ] elsewhere.

(d) /ɸ/ → [ɸ] between vowels, [p] elsewhere.

b. The phonemic representation of West Tarangan 'pig' is:

(a) /ɸɔp/

(b) /pɔp/

(c) /pɔɸ/

(d) /ɸɔɸ/

5. Sawai (Maluku, Indonesia). Sawai /d/ has two allophones: [d] and [ɾ].

1. [ndɛɾɛɾɛ] ‘he follows’
2. [mɛɾmɛɾ] ‘a plant’
3. [mdiɾkɛ] ‘you (sg.) move’
4. [deko] ‘friend’
5. [mdi] ‘tree trunk’
6. [dɛɾe] ‘wall’
7. [wɔɾɛ] ‘already’
8. [dɛl] ‘girlfriend’
9. [ndɛ] ‘he goes’
10. [nɛɾoɾ] ‘he asks’

a. Which of the following statements accounts for the distribution of these allophones.

(a) /d/ → [ɾ] between vowels, [d] elsewhere.

(b) /d/ → [ɾ] when not word-initial, [d] elsewhere.

(c) /d/ → [ɾ] after vowels, [d] elsewhere.

(d) /d/ → [ɾ] before vowels, [d] elsewhere.

b. What is the phonemic representation of Sawai ‘he follows’?

(a) /ndɛdɛɾɛ/

(b) /nɾɛɾɛɾɛ/

(c) /ndɛɾɛdɛ/

(d) /ndɛdɛdɛ/

# Chapter 4

# Morphology

## Contents


| | | |
|---|---|---|
| 4.1 | Morphemes | 72 |
| 4.2 | Morphemes and Allomorphs | 82 |
| 4.3 | Productivity | 88 |
| 4.4 | Hierarchical Structure in Morphology | 89 |
| 4.5 | The 'cranberry' problem | 92 |
| 4.6 | Summary | 93 |
| 4.7 | Exercises | 94 |


## 4.1 Morphemes

Until now, we have been concerned only with the sound system of language. We have paid little attention to things that seem like they might be somewhat important such as 'meaning.' While we have learned that English /kæts/ will have the form [kʰæts] when the grammar is finished processing it, we have done nothing to account for the fact that speakers with English-type grammars will instantly recognize that /kæts/ is a 'complex' word, in the sense that it contains more than one meaningful element. The study of the internal structure of complex words of this type is known, in linguistics, as **morphology**.

The word /kæts/, to be specific, appears to consist of two components: /kæt/, which means something like CAT, and /s/, which means something like MORE THAN ONE. Put together, as here, they therefore mean 'more than one cat.' Notice that although /kæt/ consists itself of three phonological segments, it cannot be further divided into meaningful component elements (so, e.g., it is not the case that /k/ means 'four-pawed,' /æ/ 'pet' and /t/ 'coughs up furballs'). /kæts/ has exactly two meaningful elements within it. Note, moreover, that it is not the case that any old combination of /kæt/ and /s/ gives rise to the meaning 'more than one cat': neither /skæt/, /ksæt/, nor /kæst/, all of which consist of a combination of /s/ and /kæt/, mean 'more than one cat.' *Only* /kæts/ means 'more than one cat.' Why should this be so?

It is also worthwhile to observe that if you and a friend were walking along and saw what appeared to be a very large and oddly colored hairpiece spread out on the hood of a car, and your friend said 'What's that?' you, with your masterful detective skills, might be able to answer '/kæt/!' However, if your friend asked, 'How many cats are there?' you could not answer '/s/!' to mean 'more than one!' That is, /kæt/ appears to mean CAT all on its own, while /s/ means 'more than one' *only* when attached to some noun. In particular, as we saw in the paragraph above, it must be attached to the *end* of some noun. Thus /kæts/ means 'more than one cat,' /hæts/ means 'more than one hat,' and /læps/ means 'more than one lap.'

A string of one or more phonemes which conveys a particular meaning is known as a **morpheme**. Thus /kæt/ 'cat' is a morpheme, as is /s/ 'more than one.' [Note that this does not mean that every sequence /kæt/ in English contains the morpheme /kæt/ – /skæt/, e.g., does not – any more than every /s/ in English means 'more than one.'] Moreover, /kæt/ is what is generally referred to as a **free morpheme**, meaning it can be used on its own to convey the meaning 'cat' (as in the example in the last paragraph), while the plural marker /s/ by contrast is known as a **bound morpheme**, since it must be 'bound' to another morpheme for it to convey its meaning. A bound morpheme is also frequently called an **affix**. Morphemes are the smallest units which carry meaning. Phonemes, on the other hand, are the smallest units which signal a *difference* in meaning but are not meaningful in and of themselves (they are just sounds).

Bound morphemes come in various types, depending upon where they attach to their host morpheme. Affixes which attach at the end of their host, like the English plural marker /s/ in /kæts/, are known as **suffixes**. English also has **prefixes**, which attach to the beginning of their host word. For example, /ri-/ 'again' as in /ritaj/ 'retie' (i.e., 'tie again'), or /ən-/ 'not' as in /ənbəlivəbl̩/ 'unbelievable' (i.e., 'not able to be believed'). Note that /ənbəlivəbl̩/ has both a prefix (/ən/) and a suffix (/-əbl̩/), added to the 'free morpheme' /bəliv/ 'believe'.[1]

Extremely rare in English-type grammars, indeed for some speakers non-existent, is the third logical possibility for the placement of an affix relative to its host morpheme: infixation. **Infixes** are placed somewhere inside their host morpheme. Since infixes are rare to non-existent in English, we will cite a few examples from other languages. Examine the singular : plural pairs in Oaxaca Chontal (Mexico) in the data below.

(1) Oaxaca Chontal Plurals

- /tuwa/ 'foreigner' : /tulwa/ 'foreigners'
- /lipo/ 'possum' : /lilpo/ 'possums'
- /sewiʔ/ 'magpie' : /selwiʔ/ 'magpies'
- /mekoʔ/ 'spoon' : /melkoʔ/ 'spoons'

How does one make plurals in Oaxaca Chontal? Unlike other affixes, whose placement is clear by virtue of their names, one must always specify precisely where to infix an infix. For example, if all you stated was that Oaxaca Chontal made its plurals by infixing /l/ 'more than one' and that /sewiʔ/ was the word for 'magpie,' we would not know whether the plural was /slewiʔ/, /selwiʔ/, /sewliʔ/, or /sewilʔ/! An infix is not randomly inserted anywhere within a morpheme, its placement is rule-governed. You must determine, after examining the data, exactly what the rule for placement of the infix is. (For example, if you were to say 'after the first vowel', this would work for Oaxaca Chontal).

Here's some more data on infixation, just to make it clear to you. This comes from Agta (a Phillipine language).

(2) Agta Tense Formation

[1]We use a hyphen to indicate where the affix attaches to its host. In the case of an infix, there would be hyphens on both sides.

- /gafutan/ 'grabs' : /ginafut/ 'grabbed'
- /daŋagan/ 'hears' : /dinaŋag/ 'heard'
- /hulutan/ 'follows' : /hinulut/ 'followed'
- /paligatan/ 'hits' : /pinaligat/ 'hit (past)'

Notice first that Agta has a suffix which indicates 'present tense': /an/.[2] So /gafutan/ 'grabs' is a complex word, having two parts: /gafut/ 'grab' (the root) and the suffix /an/ 'present tense.' The past tense form of this verb is created by taking the part of /gafutan/ which means 'grab' (/gafut/) and infixing into it the past tense morpheme /in/. Where is the past tense morpheme infixed in Agta? That's right, after the first consonant of the host morpheme (the verb).

It is critical that you realize that morphological surface forms are not derived from one another. That is, if you want to make the past tense of 'hit' in Agta, you do not first make the present /paligatan/, then subtract the present tense suffix /an/, then insert the past tense infix /in/. Why bother making the present tense through suffixation in this way when what you want to say is the past tense? This would be comparable in English to saying that, in order to make the past tense of the verb /glajd/, you take the 3rd singular present form /glaijdz/, subtract the 3rd singular marker /-z/ and add the past tense marker /-əd/. This seems pretty silly when stated in terms of English to you, as an English speaker, and it would seem just as silly if proposed to an Agta speaker for Agta. To say the past tense of 'hit', you take the morpheme that means 'hit' (/paligat/) and infix into it the past tense marker (/in), giving /pinaligat/. There is no reason to mess around with the present tense form.

We will offer one more example. This one comes from Katu, a Katuic language of Vietnam.

(3) Katu Infixation

- /gap/ 'cut' : /ganap/ 'scissors'

[2]As in phonology, the abstract mental representation which serves as input to the grammar is placed between slashes so both phonemic and morphemic representations are cited with slashes.

- /juut/ ‘rub’ : /januut/ ‘cloth’
- /panh/ ‘shoot’ : /pananh/ ‘crossbow’
- /piih/ ‘sweep’ : /paniih/ ‘broom’

The infix here is clearly /an/, but what does it mean? Well, ‘scissors’ are what you use to ‘cut’ with, ‘cloth’ is what you use to ‘rub’ with, a ‘crossbow’ is what you shoot with (if you are a Katu, – we will refrain from the obvious corollary), and a ‘broom’ is what you ‘sweep’ with. The nouns identify the *instrument* used to perform the action of the verb. They are therefore called ‘instrument nouns’ and the infix /an/ is used to derive instrument nouns from verbs. As in the Agta case above, this infix is placed after the initial consonant of the verb which hosts the infix.

There is one more kind of morpheme. This is usually referred to as a ‘process morpheme’ since, rather than consisting of some string of phonemes, it requires, instead, modification of the morpheme to which it attaches. We will consider only one type of process morpheme here: reduplication. Examine the data from Northern Tepehuán below.

(4) Northern Tepehuán Reduplication

- /toʃi/ ‘rabbit’ : /totoʃi/ ‘rabbits’
- /kəli/ ‘man’ : /kəkəli/ ‘men’
- /mara/ ‘son’ : /mamara/ ‘sons’
- /tova/ ‘turkey’ : /totova/ ‘turkeys’
- /ʃiəgi/ ‘older brother’ : /ʃiʃiəgi/ ‘older brothers’
- /sukuli/ ‘younger brother’ : /susukuli/ ‘younger brothers’
- /dəgi/ ‘rat’ : /dədəgi/ ‘rats’

Can you figure out how Northern Tepehuán forms the plural of these words? Of course you can. One takes the initial consonant and vowel of the base morpheme and prefixes them (in their original order) to that base. This process, doubling some of the material already present in the morpheme to which the process applies, is known as 'reduplication.' The meaning of reduplication in these Northern Tepehuán words is 'more than one.'

Below is another case of reduplication. (It's a popular process . . . ) This one comes from our old friend, Agta.

(5) Agta Reduplication

- /adanuk/ 'long' : /adadanuk/ 'very long'
- /addu/ 'many' : /adaddu/ 'very many'
- /apisi/ 'small' : /apapisi/ 'very small'
- /abikan/ 'near' : /ababikan/ 'very near'

Here, the initial vowel plus consonant sequence of these words is being reduplicated, the doubled segments being prefixed in their original order to the base morpheme. The meaning is *intensification* (i.e., an extreme version of whatever the base morpheme means).

It is necessary, at this point, to add a couple of cautionary statements, the first about how one should regard data generally and the second about bound morphemes (in particular, infixes). We will use the Southern Barasano below to illustrate our points.

(6) Southern Barasano [A vowel with a tilde over it is nasalized, as was stated earlier.]

1. /wabõ/ 'she goes'
2. /bĩdĩbĩ/ 'he goes upstream'
3. /tıdibı/ 'I returned' and 'you return'
4. /bĩdĩkoabĩ/ 'he really goes upstream'
5. /tıdiboabı/ 'I unexpectedly return' or 'you unexpectedly return'
6. /wakoabı/ 'I really go' or 'you really go'
7. /tıdirũtũbĩ/ 'he continues to return'
8. /bĩdĩboabõ/ 'she unexpectedly goes upstream'

9. /warũtũbĩ/ 'he continues going'
10. /tɪdikoabõ/ 'she really returns'
11. /bĩdĩrũtũbɪ/ 'I continue to go upstream' or 'you continue to go upstream'
12. /waboabõ/ 'she unexpectedly goes'

This appears rather complicated. The first thing to realize regarding data in general (in case it is not already obvious) is that the morphological entities found in one language do not necessarily, or even usually, correspond in a nice one-to-one way to those found in another language.[3] Thus, while intensification is done through reduplication in Agta, in English-type grammars we usually use a free morpheme ('very') to express the same meaning. So, although we find all kinds of things in the English translation of the Southern Barasano words above, it is clear that all of the meanings contained within the English – including those many meanings which are expressed by free morphemes (words) in the English – are expressed by various types of affixes in Southern Barasano. To solve such a problem, it is necessary to identify which parts of the Southern Barasano express which meanings. The meanings will not necessary match each individual *word* of the English – sometimes English uses multiple words to express what is expressed in Southern Barasano by a single affix.

Let us first identify the Southern Barasano rendering of the pronouns of the English. These include 'he', 'she' and, rather weirdly (given the languages of the world), something that apparently means both 'I' and 'you' (since every sentence involving 'I' in the data above can also be translated 'you', and vice-versa). If you look at all of the sentences involving 'he' in the English translation, you will see that they each end in /bĩ/. Apparently the suffix /bĩ/ means 'he'. If you look at those sentences which involve 'I' and 'you', you will see similarly that the suffix /bɪ/ occurs in all Southern Barasano words which are translated with 'I' and 'you' in English. How about 'she'? Well, it is a little hard to tell without further analysis whether 'she' is /bõ/ or /abõ/ – since all 'she' sentences end with the same last three segments in Southern Barasano. We will keep the options open for 'she' for the moment and pass on to identify what in the Southern Barasano corresponds to English 'really'.

[3]This is just as true in phonology and syntax as it is in morphology.

If you look at the relevant examples, you will quickly see that it is /koa/. So 'I/you really go' in (6) appears to consist of /wa/ 'go' plus /koa/ 'really' plus /bɪ/ 'I/you' (thus meaning 'I/you really go').[4] That means that 'she goes' in (1) is probably /wa/ 'go' plus /bõ/ 'she.' Now we have gotten 'she'. How does Southern Barasano express the fact that the action indicated by the verb happened unexpectedly? Clearly, /boa/ is responsible for this part of the meaning of the Southern Barasano words above. And how does Southern Barasano express that the action indicated by the verb is continuing to go on? – with /rũtũ/. This leaves the verbs (which appear to be morphemes from the data) /wa/ 'go', /bĩdɪ̃/ 'go upstream',[5] and /tɪdi/ 'return'.

Now comes the second caution. Given that the analysis of (9) is /wa/ 'go' plus /rũtũ/ 'continue' plus /bɪ̃/ 'he' (therefore meaning 'he continues going'), is /rũtũ/ an infix? The answer is a resounding **NO**. An infix is inserted *into* a morpheme. /rũtũ/ is not inside some other morpheme – it's merely been suffixed to the verb /wa/ *before* /bɪ̃/ was suffixed. That is, we must distinguish between a series of suffixes (or prefixes) and a true infix. To be an infix, you must be *inside* another morpheme, not merely *between* two other morphemes. Do not confuse infixes with suffixes or prefixes. In addition, when we find multiple suffixes, as here, we must state, when we give a solution to the problem, what order the suffixes attach in. (This obviously applies to prefixes, as well.)

The simplest type of morphological problem consists merely of *segmenting* complex words into their component parts. Just for practice, we will give three instances of this type of problem here. The answer will be on the next page. The task, in case it is not clear, is to identify the morphemes, state whether they are bound or free, and, if bound, whether they are prefixes, suffixes, infixes, or reduplication (a process morpheme). If there are multiple prefixes or suffixes, state the order.

(7) Michoacán Nahuatl

- /nokali/ 'my house'
- /nokalimes/ 'my houses'

[4] There appears to be nothing here indicating 'present tense' – this is not unusual in the languages of the world.

[5] Note that in Southern Barasano, unlike in English, 'go upstream' appears to be a unitary morpheme – i.e., it does not contain 'go'. Again, this kind of mismatch between languages is common.

- /mokali/ 'your house'
- /ikali/ 'his house'
- /kali/ 'house'
- /nopelo/ 'my dog'
- /mopelo/ 'your dog'
- /mopelomes/ 'your dogs'
- /pelo/ 'dog'
- /nokwahmili/ 'my cornfield'
- /ikwahmilimes/ 'his cornfields'
- /kwahmili/ 'cornfield'

(8) Isthmus Zapotec

- /ʒigi/ 'chin'
- /kaʒigi/ 'chins'
- /ʒigibe/ 'his chin'
- /ʒigiluʔ/ 'your (sg.) chin'
- /kaʒigitu/ 'your (pl.) chins'
- /diaga/ 'ear'
- /kadiagatu/ 'your (pl.) ears'
- /kadiagadu/ 'our ears'
- /biʃozedu/ 'our father'
- /biʃozetu/ 'your (pl.) father'
- /kabiʃozetu/ 'your (pl.) fathers'

(9) Pocomchí [/q/ is a voiceless, uvular stop. Can you find your uvula?]

- /suk/ ‘good’ : /suksuk/ ‘delicious’
- /raʃ/ ‘green’ : /raʃraʃ/ ‘very green’
- /nim/ ‘black’ : /nimnim/ ‘jet black’
- /kaq/ ‘red’ : /kaqkaq/ ‘very red’
- /saq/ ‘white’ : /saqsaq/ ‘very white’

The Answer to the Michoacán Nahuatl Problem

There are three free morphemes in the data: /kali/ ‘house’; /pelo/ ‘dog’; and /kwahmili/ ‘cornfield’.

There is one suffix in the data: /mes/ ‘plural’.

There are three prefixes in the data: /no/ ‘my’; /mo/ ‘your’; and /i/ ‘his’.

The Answer to the Isthmus Zapotec Problem

There are three free morphemes in the data: /ʒigi/ ‘chin’; /diaga/ ‘ear’; /biʃoze/ ‘father’[6]

There are four suffixes in the data: /be/ ‘his’; /luʔ/ ‘your (sg.)’;[7] /tu/ ‘your (pl.)’; and /du/ ‘our’.

There is one prefix in the data: /ka/ ‘plural’.

The Answer to the Pocomchí Problem

[6]Note that although ‘father’ is not found standing on its own in the data given, the most general solution to the problem – and thus the most valuable – is that which can see beyond the limited data given. Since all of the other nouns which show up with the various prefixes and suffixes present in the data can also stand on their own as ‘free morphemes,’ it is logical to conclude that the same is true of ‘father’.

[7]Isthmus Zapotec, like many human languages, distinguishes between a singular ‘you’ and a plural one. English used to do this with ‘thou’ vs. ‘ye’.

The free morphemes are all those listed before the colon. The other morpheme is complete (or ‘full’) reduplication, which indicates *intensification*. As we have now seen, reduplication may be *partial* (cf. Agta) or *full* as in Pocomchí.

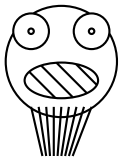
## 4.2 Morphemes and Allomorphs

It would be nice if morphological life were as simple as what we have seen so far – unfortunately, it is not. The English plural forms below illustrate some further complexities.

(10) English Plurals

- /kæt/ ‘cat’ : /kæts/ ‘cats’
- /rɪp/ ‘lip’ : /rɪps/ ‘lips’
- /bæk/ ‘back’ : /bæks/ ‘backs’
- /klɪf/ ‘cliff’ : /klɪfs/ ‘cliffs’
- /kæd/ ‘cad’ : /kædz/ ‘cads’
- /rɪb/ ‘rib’ : /rɪbz/ ‘ribs’
- /bæg/ ‘bag’ : /bægz/ ‘bags’
- /hajv/ ‘hive’ : /hajvz/ ‘hives’
- /bənænə/ ‘banana’ : /bənænəz/ ‘bananas’

From this list, you can see quite clearly that whereas some nouns make their plural by suffixing /s/, others do so by suffixing /z/. While it is possible that we all just memorize which nouns take /s/ and which take /z/, there is considerable evidence that this is not correct. In order to see why, examine the figure below.



Suppose we tell you that this is a picture of the dreaded /birdəd slorg/. We can safely assume that you have never heard of a /birdəd slorg/ before since we just created it afresh. If we were to ask you, now, what the following is a picture of:

you would invariably give the answer that this is a picture of two /birdəd slorgz/. Psycholinguistic experiments have shown that this is precisely what speakers with English-type grammars answer in such a testing situation. The interesting question is this: why do all people with English-type grammars answer this question in the same way? You cannot have *known* what the plural of /slorg/ was before you were asked, since you had never heard the word before, and we did not tell you the plural.

Notice that if we now told you that we had made an error in judgement and that in fact the first picture above was not a /birdəd slorg/ at all, but rather the benign /birdəd zorp/, and extracted from you what the second picture contained, you would say it showed 'two /birdəd zorps/.' Again, this has been conclusively demonstrated in experiments performed on willing subjects.

If you examine the list of English plurals given in the table in (10) above, the reason for this otherwise inexplicable behavior should be clear. English regular plurals (i.e., ignoring things like 'sheep' : 'sheep', 'ox' : 'oxen', 'goose' : 'geese' – to which we will turn our attention momentarily) appear to add /s/ after voiceless segments, but /z/ after voiced ones. Now consider the data below.

(11) More English Plurals

- /mɛs/ 'mess' : /mɛsəz/ 'messes'

- /mez/ ‘maze’ : /mezəz/ ‘mazes’
- /ræʃ/ ‘rash’ : /ræʃəz/ ‘rashes’
- /gəraʒ/ ‘garage’ : /gəraʒəz/ ‘garages’
- /mæʧ/ ‘match’ : /mæʧəz/ ‘matches’
- /bæʤ/ ‘badge’ : /bæʤəz/ ‘badges’

These examples show that if a noun ends, in English, in an alveolar fricative or an affricate, it takes the plural marker /əz/ (regardless of the voiced or voiceless status of its final segment).

For regular plurals in English, the precise form of the plural marker (/s/, /z/, or /əz/) is fully predictable. In keeping with our general practice of not admitting predictable information into the lexicon, we therefore want to store only *one* of these forms. The others can be derived by a rule which produces the variant realizations of this one underlying morpheme.[8] The variant realizations of a morpheme are called *allomorphs*, much as the variant realization of a phoneme are known as allophones. The regular plural marker in English therefore has three allomorphs: /s/, /z/, and /əz/. The rules governing these forms are stated informally below.

- Rules for English Regular Plural Formation (informal statement, to be revised shortly)
    1. If the noun ends in an alveolar fricative or an affricate, add /əz/.
    2. If the noun ends in a voiceless consonant which does not fall under Rule 1, add /s/.
    3. If the noun ends in a voiced consonant which does not fall under Rule 1, add /z/.

When we were dealing with allophones above, it became clear that when we had, for example, two allophones of a given phoneme ([p] and [pʰ] as

[8]In addition, we should note that speakers, if asked, appear to be unaware that they produce three different forms when they produce plurals. This lends further support for the fact that only one of the forms needs to be stored (explicitly memorized) and the others are produced automatically by rule.

allophones of the phoneme /p/, for instance) we needed *one* rule to produce them. This is because if we posit a rule which aspirates voiceless stops (thus producing, under the proper conditions, [$\text{p}^\text{h}$] from the phoneme /p/) the other allophone ([p]) would surface by default if we simply applied no rule at all to it. The same is true for allomorphs. We need to pick one of the regular plural allomorphs (/s/, /z/, or /əz/) as the underlying morpheme, then by positing *two* rules, we can produce two allomorphs. The third allomorph will be the result of doing nothing at all to the underlying morpheme, and thus does not require a rule at all.

Now imagine that we were to select /əz/ as the underlying form of the regular plural morpheme in English. We would need (1) a rule which deleted the schwa (/ə/) except after an alveolar fricative or an affricate, (2) a rule which forced the /z/ that resulted from the first rule to *assimilate* (i.e., adopt the same value for some feature, in this case voicing) to the voicelessness of a final voiceless consonant, and (3) a statement that the first rule must be applied before the second.

If we assumed /s/ was the underlying form for the regular plural morpheme, we would need (1) a rule which inserted schwa before this /s/ if the noun to which the plural marker was attached ended in an alveolar fricative or an affricate, (2) a rule that assimilated /s/ to the voicing of the segment immediately preceding it, and (3) a statement that the first rule applies before the second (since the forms affected by the first rule need to come out /əz/, not /əs/). Note that the second rule required under this assumption cannot be a general *phonological* rule of English, since the final 's' of /fɔls/ 'false' does not undergo it (i.e., /fɔls/ is not pronounced [fɔlz]).[9]

Finally, if we assumed /z/ was the underlying form for the regular plural morpheme, we would need (1) a rule which inserted schwa before this /z/ if the noun to which the plural marker was attached ended in an alveolar fricative or an affricate, (2) a rule that assimilated /z/ to the voicelessness of the segment immediately preceding it, and (3) a statement that the first rule must apply before the second.

The second of these options (in which the regular plural morpheme is /s/) is the least satisfactory for the following reason. It is more economical to have a system with less exceptions than more exceptions. Because phonological rules operate without exception, a solution which relies on phonological rules

[9]One characteristic of phonological rules, as opposed to morphological rules, is that they operate 'across the board' – there are no exceptions to phonological rules.

is preferred. It is not always possible to have a purely phonological solution to a morphology problem, but when it *is* possible, that option should be chosen. One of the rules which choosing /s/ as the morpheme requires does not appear to be generally true of English, given the existence of words such as /fɔls/, thus, you cannot have a phonological solution. This problem does not plague either other solution. So, how do we pick between the two remaining options? Is it /z/ or /əz/? The key here is in the way in which the first rule is stated for each of these options. In the case of the /əz/ choice, the rule says 'delete schwa *except* after an alveolar fricative or an affricate'. In the case of the /z/ choice, the rule says 'insert a schwa after an alveolar fricative or an affricate'. Casually termed, for the first choice, /əz/, we need a rule which says 'don't delete schwa here' (where 'here' is after an alveolar fricative or affricate). For the second choice /z/, we need a rule that says 'do insert schwa here' (same 'here' as above). It is a general principle of grammar, for reasons of economy such as those stated earlier, that rules are stated in positive terms with well-defined environments. Therefore, 'do insert schwa here' (where 'here' is the well-defined environment of 'after alveolar fricatives and affricates') is the preferred choice. The final choice, then, for the regular plural morpheme in English-type grammars is /z/.

- Regular Plural Allomorph Creation Rule (informal, final version)

    1. Insert a schwa if the plural morpheme follows an alveolar fricative or an affricate.
    2. $z \rightarrow s$ after voiceless consonants

The rules apply in the order stated. A similar rule accounts for the regular third person singular present tense verb forms in /z/ (with allomorphs /z/, /s/, and /əz/ as in /goz/ 'goes', /kɪks/ 'kicks' and /wɪʃəz/ 'wishes') and the possessive.

The English plural is a fairly complex case so let us look at something simpler. Examine the Sierra Popoluca data below.[10]

1. tək 'house' : antək 'my house'

[10]We will omit brackets of any type from now on unless they are relevant. It should be reasonably clear what level, phonetic or phonemic, we are dealing with at any particular time.

2. petkuj 'broom' : ampetkuj 'my broom'
3. meːme 'butterfly' : ammeːme 'my butterfly'
4. heːpe 'cup' : anheːpe 'my cup'
5. kawah 'horse' : aŋkawah 'my horse'
6. nəːyi 'name' : annəːyi 'my name'
7. suuŋ 'cooking pot' : ansuuŋ 'my cooking pot'

What kind of affix expresses possession by the speaker in Sierra Popoluca? A prefix. What are its allomorphs? Well, the affix shows up in the following forms in the data above: /an/ (1,4,6,7), /am/ (2,3), and /aŋ/ (5). 'Is the distribution predictable?' is obviously the next question. (If you simply keep on reading to find out the answer, you will never learn to solve problems of this type on your own. Take a minute to look at the actual data.) Yes, the distribution is predictable. /am/ occurs before /p/ and /m/, /aŋ/ occurs before /k/, and /an/ occurs before /t/, /h/, /n/, and /s/. Is there a reason /am/ might occur before /p/ and /m/? Yes, indeed. /p/ and /m/ are both *labial* and the nasal at the end of the allomorph of 'my' that occurs before words that start with these consonants is *labial*, too. Does it make sense that we get the allomorph /aŋ/ before /k/? Yes it does. /k/ is a *velar*, and /aŋ/ is the variant which has a velar nasal. Does is make sense that we get /an/ before /t/, /h/, /n/, and /s/. Well, partly. /t/, /n/, and /s/ are all alveolars, as is the /n/ at the end of this allomorph for 'my', but what about /h/. Is /h/ alveolar? Of course not. So when we add 'my' to a noun in Sierra Popoluca the final nasal of that morpheme always matches the initial consonant of the noun in place of articulation, except in the case of /h/. What that tells us that the form before /h/ must be the *default* form of the morpheme – i.e., the form that surfaces when nothing special happens to it. So we should assume that /an/ is the underlying form of 'my' and that it assimilates in place of articulation to following labial and velar consonants. Thus underlying /an/ will become /am/ before labials and /aŋ/ before velars. It won't do anything before /s/ or /h/ – and thus will come out unchanged (i.e., as its underlying form, /an/). This is clearly the correct result.

## 4.3 Productivity

The case of the English plural, which we discussed above, illustrates something else quite important about morphology. Depending upon one's theory of morphology, one might say that there are two kinds of morphemes: productive and unproductive. A **productive** morpheme can be freely used with 'new' lexical items, such as the word 'slorg' which we invented in the previous section. Any native speaker of English can generate the plural of a noun such as 'slorg' (e.g., if we had made up 'blorg,' 'norg', 'lorg', or 'horg', everyone who has an English-type grammar would agree that their plurals are /blorgz/, /norgz/, /lorgz/, and /horgz/). Another example is the English nominal prefix 'ex-' (meaning 'former'). If we shaved off the beard of the 'slorg', it could be an 'ex-slorg.' Similarly, given the other nouns we just made up, we can make 'exblorg,' 'exnorg,' 'exlorg,' and 'exhorg.' The sense of 'he's an exblorg' would immediately be clear – he used to be a blorg, but isn't one anymore. Note that this all works even if you do not know exactly what a 'blorg' is (which is presumably the case). The regular plural and 'ex-' are both productive morphemes.

We can contrast this with an **unproductive** morpheme such as the suffix in 'warmth', which clearly consists of 'warm,' meaning something like 'warm,' and /-θ/, meaning 'the property designated by the adjective to which I am suffixed.' Thus 'warmth' means 'the property designated by the adjective 'warm'. We find the same suffix in 'strength' (from 'strong' – note the wacky vowel change), 'length' (from 'long'), 'depth' (from 'deep'), 'width' (from 'wide'), 'filth' (from 'foul'!!), etc. The first thing you notice is that, unlike, for example, '-ness,' which does a similar job, you cannot make words like *'uglith' (the abstract property shared by all ugly things), *'coolth' (or *'cilth' or *'celth' or whatever, meaning the abstract property shared by all 'cool' things), or *'hetth' (the abstract property shared by all 'hot' things – with the vowel change seen in 'long':'length' and 'strong':'strength'). The abstract nouns made by /-θ/ are listable and finite – no new ones can be created (unlike plural and 'ex-' forms). 'Ugliness,' 'coolness,' and 'hotness' are all fine, on the other hand. In fact, '-ness' is quite productive. If we make up a new adjective, like 'slarm' meaning 'both chunky and bearded', we can easily say that someone is richly endowed with 'slarmness', but *'slarmth' is not going to be understood by anyone, even someone who you just told what 'slarm' means.

How are we to explain the difference between productive and unproductive morphology? They key resides in the fact that while it is possible, even trivial, to list all abstracts nouns in /-θ/, it is *not* possible to list all plurals in /-z/. Any newly formed word will make its plural in /-z/, but no newly-created words will make abstracts in /-θ/. How are we to capture this fact in the model of linguistic knowledge we have been working on? Listable, non-predictable information, such as the fact that 'strong' has an abstract 'strength,'[11] must be stored in the lexicon, just like unpredictable phonological information is stored there (e.g., the fact that 'bean' starts with a voiced bilabial stop). On the other hand, if a noun does not have an irregular (and thus unpredictable) plural stored with it (like 'sheep,' plural 'sheep'), its plural is predictable: it will be in /-z/ (which may be articulated as [s], [z], or [əz]). It would be silly to list all these regular plurals in the lexicon – every noun which is not irregular in the plural has this regular plural. So the lexical entry for 'dog' will not contain *any* information about its plural form – then the default, predictable /-z/ plural will be used with this noun. On the other hand, 'sheep' will contain the information in its lexical entry that says, 'Oh, by the way, my plural is 'sheep.'

Now we see why newly coined words like 'slorg' make a plural in /-z/. When we told you that there is a word 'slorg', we did not tell you that it had any kind of irregular plural. Therefore, you stored it in your lexicon with no information about how its plural is formed. This is, of course, just like the way 'dog' or 'banana' is stored. All such nouns show the plural morpheme in /-z/. Productive morphology is simply predictable, default morphology. Non-productive morphology is unpredictable, irregular morphology.

## 4.4 Hierarchical Structure in Morphology

So far, we have looked at complex words as if the type of affixation and order of affixes was all that was involved – this was something of a simplification. Consider the complex word /ən-taj-əbl̩/. We have used to dashes to separate this word into its component morphemes. Note, however, that the word itself is **ambiguous**, i.e., has more than one interpretation. On the one hand, it

[11] Note that as a speaker with an English-type grammar, this fact is 'predictable' by you in some sense. But that's cheating – it's predictable precisely because you already know that it is the case. The question of predictability is one of whether or not someone who isn't cheating – by already knowing the answer – can predict whether 'strong' makes 'strength' or 'strongness.' They cannot.

means 'able to be untied,' as in: 'that knot *is* untieable, I untied it in just five minutes!' On the other hand, it can also mean 'not able to be tied,' as in 'that knot is definitely untieable, I've been trying to tie it for twenty years and I'm a world-renowned master at tying knots!' How are we to account for the fact that this one word has two very different interpretations?

Let us examine the matter of ambiguity in more detail. Ambiguity comes in two well-defined types: **lexical** and **structural** ambiguity. **Lexical ambiguity** is based on the fact that often different meanings are associated with an identical phonological form in the lexicon. (These are **homophones**.) For example, the phonological form /najt/ is associated with a meaning which refers to a period of darkness that takes place every 24 hours (on Earth), as well as with a meaning which refers to a medieval kind of guy who rides around on a horse in armor or whatever. If I say, 'the /najt/ was very dark' you don't know whether I've just made a claim about the part of the day after sunset or about some swarthy guy on a horse. Or, take a look at the following example:

(12) She can't /ber/ children.

This sentence has at least the following meanings, based on the many different English words which have the same phonological form (/ber/). (1) She can't stand children. (2) She can't give birth to children. (3) She can't carry children ('she can't /ber/ the children across the water – they're too heavy'). (4) She can't make children naked (it's too embarrassing).[12] All of these various interpretations are based on the simple fact that the phonological sequence /ber/ is associated with numerous, disparate meanings – in this case, in particular, in several distinct verbs. It is thus a case of **lexical ambiguity**.

**Structural ambiguity**, by contrast, centers around the fact that the same morphemes may be combined in more than one order, with the interpretation matching the order of affixation. The 'un-tie-able' example is a case of structural ambiguity. The two interpretations are derived as follows:

(13) a. combine /ən/ and /taj/ to get /əntaj/ 'untie', add /-əbl̩/ giving 'able to be [ untie ] -d'

[12] For some speakers, there is yet another verb /ber/ – it means 'to maul like a bear' – so our example sentence could also mean, for such speakers, 'she can't maul children like a bear (her teeth aren't sharp enough).'

b. combine /taj/ and /-əbl̩/ to get /tajəbl̩/ 'able to be tied', add /ən-/ giving 'not [ able to be tied ]'

/-əbl̩/ in general means 'abled to be Xed' (so 'kickable' = 'able to be kicked,' 'fakable' = 'able to be faked,' etc.), so when it gets added to /əntaj/ in (13a) it turns 'untie' into 'able to be untied.' In (13b) we have added /-əbl̩/ to 'tie' directly, giving 'able to be tied.' Then, we negated this with /ən-/, giving 'not able to be tied.'

Unfortunately, in the morphology we've been talking about thus far, there is no way to represent the *order* of affixation. Put another way, the fact that 'un-' and 'tie' are closer to one another (giving 'untie', to which '-able' is added) in the first example, whereas 'tie' and '-able' are closer to one another (giving 'tieable' to which 'un-' is added) in the second is not captured by the morphological analysis that says that 'untieable' is 'un-tie-able.' It is traditional in modern linguistics to show the order/closeness facts by means of **tree** structures, as in (14) and (15).

(14) Tree for (13a)

```
        Adj
       /   \
      V    -əbl̩
     / \
   ən-  V
        taj
```

(15) Tree for (13b)

```
      Adj
     /   \
   ən-   Adj
         /  \
        V   -əbl̩
        taj
```

Let us examine these trees in some detail, because you will see similar things later on in the chapter on syntax. In the first tree (14), the fact that, in the interpretation given in (13a), 'ən-' and 'tie' are closer together is captured by graphically designating them as sisters in the tree (sisters are nodes of the tree which have a common mother node – in (14) 'un-' and 'tie' are daughters of the node labelled 'V'). The label 'V' indicates that when 'un-' attaches to a verb like 'tie' the result is still a verb (i.e., 'un-' prefixation does not change the part of speech of the morpheme to which it attaches). The little subtree over 'un-' and 'tie' is itself a sister to '-able' – indicating that '-able' is further away from 'tie' than is 'un-' (it's an 'aunt' of 'tie', if you will – rather than a sister; your sister is a closer relative than your aunt, right?). In (15) by contrast, '-able' is a sister of 'tie'. When '-able' gets added to a verb (like 'tie'), the result is not a verb, but rather an adjective – this is indicated by the fact that the mother of 'tie' and '-able' is labelled 'Adj' for 'adjective.' In this tree, 'tie' and '-able' are sisters, and thus closer to each other than either is to 'un-.' By contrast, 'un-' is a sister of the complex adjective 'tie-able.' Since 'un-' doesn't change the part of speech of an element that it attaches to, the result of adding 'un-' to an adjective is still an adjective. This is indicated by the label 'Adj' at the top of the tree.

Trees encode what we usually call 'hierarchical' information about the structure of complex words which is lacking in a simple representation such as 'un-tie-able.' Since that information is critical to the interpretation of the word, it must be part of its representation.

Finally, we can see that 'structural ambiguity' results when there are two (or more) ways in which multiple morphemes can combine into a complex morpheme (and, consequently, two or more tree structures). Since you cannot, without further information, deduce which of the two possible structures is actually in the mind of the person talking to you if someone just comes up and asks you what 'untieable' means, the string of morphemes is ambiguous. You can assign it both or either of the possible, hierarchical analyses. We will return to this matter when we discuss syntax.

## 4.5 The 'cranberry' problem

In general, morphological segmentation is quite straightforward, given sufficient evidence. It does not take a lot of effort to figure out from 'untie,' 'unwrap,' 'uncover,' and 'unbutton' that English has a verbal prefix 'un-' which means 'to reverse the action designated by the verb' (so 'unwrap' is to reverse the action of wrapping, etc.). However, consider the following examples: 'lukewarm,' 'huckleberry,' 'cranberry', and 'hamburger.' It is clear enough that the 'warm' part of 'lukewarm' means 'warm,' but what does the 'luke' part mean? Notice the weirdness of *'lukecool,' *'lukesnowy,' and *'lukesmart' (an asterisk is used in linguistics to indicate ill-formed structures). Similarly, the 'berry' part of 'huckleberry' and 'cranberry' appears to mean 'berry,' but what on earth is a 'huckle' or a 'cran'? And 'hamburger' is rather worse. A 'fishburger' is a sandwich made of fish, a 'chickenburger' is one made of chicken but a 'hamburger' is not one made of 'ham' but rather one made of ground beef. A 'cheeseburger' is not a cheese sandwich, but instead something made of ground beef with cheese on it. Does the 'ham' in 'hamburger' mean 'ground beef'? Does the 'cheese' in 'cheeseburger' mean 'ground beef with cheese on it' instead of meaning just 'cheese'? The whole affair is assuming nightmare proportions. Because the first widely discussed example of such a phenomenon – the partial analyzability of what appear to be complex words – involved 'cranberry,' these sorts of examples have come to be called 'cranberry morphemes.' Ocean Spray has compromised this terminology by the invention of such beverages (with their corresponding neologisms) as 'Cranapple,' and 'Crangrape.' They have not, however, solved the basic linguistic problem which is – are these **simple** words (consist only of a single morpheme) or **complex** words (consist of more than one morpheme)? If they are simple, then we have to somehow force ourselves to ignore those sequences of sounds which are meaningful, such as 'berry.' If, on the other hand, they are complex, we must find some meaning to assign to sequences like 'cran' since every morpheme, by definition, must have a meaning assigned to it. This is a difficult question which has been, at various times, hotly debated in the literature. While we have no doubt that you find the topic totally engaging as well, we have decided it is of less central importance than many other topics and will not pursue it here.

## 4.6 Summary

This chapter illustrates the ways in which the grammar stores and manipulates meaningful units or *morphemes*. We see that, as in phonology, there are two levels of representation with a computational component interfacing between the two. For morphology, the level of underlying representation is the *morphemic* level and the less abstract (surface) level consists of *allomorphs*. Morphological computations create *complex* words from *simple* words through systematic processes for combinining *free* and *bound* morphemes.

## 4.7 Exercises

1. Southern Barasano

   1. kahea ‘eye’
   2. kahe ‘eyes’
   3. bitia ‘bead’
   4. biti ‘beads’
   5. kĩa ‘cassava’
   6. kĩ ‘cassavas’

Which of the following statements is true about Southern Barasano?

(a) Southern Barasano has a prefix /a/ which marks the singular.

(b) Southern Barasano has a suffix /a/ which marks the singular.

(c) Southern Barasano has a prefix /a/ which marks the plural.

(d) Southern Barasano has a suffix /a/ which marks the plural.

2. Bahnar – note that the symbol ɲ is used to indicate a palatal nasal (as in Spanish <ñ>, Italian <gn>, or some English-speakers’ pronunciation of ‘onion’ [əɲjən]).

   1. həɲír ‘sweaty person’
   2. təbúl ‘person to sleep crawled up with’
   3. məlwɛ́ŋ ‘drunk person’
   4. həháh ‘laughing and giggling girl’
   5. həɲír həɲóɲ móɲ ‘each of many sweaty people’
   6. məlwɛ́ŋ məlwóɲ móɲ ‘each of many drunk people’

7. həháh həhóɲ móɲ ‘each of many laughing and giggling girls’

a. What morphological process do we see at work in the Bahnar data?

(a) replication

(b) reconstruction

(c) reduplication

(d) stuttering

b. How would you say ‘each of many people to sleep curled up with’ in Bahnar?

(a) təbúl təbúl móɲ

(b) təbúl təbóɲ móɲ

(c) təbóɲ təbúl móɲ

(d) jəbədəbədu

3. Xavante

- du ‘stomach’
- ʔaddu ‘your stomach’
- ʔra ‘child’
- ʔajʔra ‘your child’
- hiʔrãti ‘knee’
- ʔajhiʔrãti ‘your knee’
- tɔ ‘eyes’
- ʔattɔ ‘your eyes’
- ʔwa ‘tooth’

- ʔajʔwa ‘your tooth’
- brɔ̃ ‘wife’
- ʔajbrɔ̃ ‘your wife’
- ʃɛːrɛ ‘hair’
- ʔaʃʃɛːrɛ ‘your hair’
- paːra ‘foot’
- ʔajpaːra ‘your foot’
- bãːbã ‘father’
- ʔajbãːbã ‘your father’

a. What are the allomorphs of the morpheme meaning ‘your’ in Xavante?

(a) ʔaj, ʔat, ʔad, and ʔaʃ

(b) ʔaj and ʔaʃ

(c) ʔa

(d) jəbədəbədu

b. What is the underlying morpheme meaning ‘your’ in Xavante?

(a) ʔaj, ʔat, ʔad, and ʔaʃ

(b) ʔaʃ

(c) ʔad

(d) ʔaj

c. The precise rule responsible for the allomorphs of ‘your’ in Xavante is:

(a) /j/ assimilates completely to (i.e., becomes the same as) following stops

(b) /j/ assimilates completely to (i.e., becomes the same as) following alveolars and alveopalatals

(c) /j/ assimilates completely to (i.e., becomes the same as) following voiceless consonants

(d) /j/ assimilates completely to (i.e., becomes the same as) following stops and fricatives

4. Waorani

1. abo 'I see'
2. adã 'she sees'
3. amõ 'we (inclusive) see'
4. ãmo 'I say'
5. ãŋã 'he says'
6. kǣnã 'she eats'
7. kǣmõ 'we (inclusive) eat'
8. kækã 'he does'
9. kæmõ 'we (inclusive) do'

a. The allomorph(s) of 'I' in Waorani is or are:

(a) /bo/ and /mo/

(b) /mo/

(c) /bo/

(d) /abo/

b. The allomorph(s) of 'she' in Waorani is or are:

(a) /da/

(b) /dã/ and /nã/

(c) /da/ and /na/

(d) /na/

c. The allomorph(s) of 'he' in Waorani is or are:

(a) /kã/ and /ŋã/

(b) /ka/ and /ŋa/

(c) /a/ and /ã/

(d) /ŋã/

d. The rule governing the distribution of all of these allomorphs in Waorani is:

(a) nasalize stops after non-nasal vowels

(b) denasalize stops after non-nasal vowels

(c) nasalize stops after nasal vowels

(d) denasalize stops after nasal vowels

e. If you were about to marry a Waorani person and were asked during the marriage ceremony to assent to the marriage, you would probably say:

(a) kæmo

(b) kæbo

(c) kæ̃mo

(d) hə

5. Copainalá Zoque

1. kenu ‘he looked’
2. kenpa ‘he looks’
3. sihku ‘he laughed’
4. sikpa ‘he laughs’
5. wihtu ‘he walked’
6. witpa ‘he walks’
7. kaʔu ‘he died’
8. kaʔpa ‘he dies’
9. nahpu ‘he kicked’
10. nahpa ‘he kicks’
11. sohsu ‘he cooked’
12. sospa ‘he cooks’

a. The morpheme which marks the past tense in this variety of Zoque is:

(a) the suffix /u/
(b) the prefix /u/
(c) the infux /u/
(d) reduplication

b. The morpheme which marks the present tense in this variety of Zoque is:

(a) the suffix /a/
(b) the suffix /pa/
(c) the infix /p/
(d) the infix /a/

c. There is some variation in the verbal roots involving /h/. What rule triggers this variation?

(a) delete /h/ if it is followed by a stop

(b) delete /h/ if it is followed by two consonants

(c) delete /h/ if it is not followed by two stops

(d) delete /h/ whenever you don't feel like saying it

d. What is going on with Zoque 'kick'?

(a) The root is /nah/ and the present tense has an irregular /p/.

(b) The root is /nahp/ and there is a rule which deletes one of two adjacent /p/'s.

(c) The root is /nahp/ and there is a rule which deletes /p/ after /h/.

(d) The root is /kɪk/ and nothing's going on with it.

e. Do the rules of /p/ deletion and /h/ deletion need to be ordered with respect to one another?

(a) No.

(b) Yes. /p/ deletion needs to occur after /h/ deletion.

(c) Yes. /h/ deletion needs to occur after /p/ deletion.

(d) Yes, but it does not matter which order you use.

# Chapter 5

# Syntax

## Contents


5.1 Introductory remarks . . . . . . . . . . . . . . . . 101
5.2 'Grammaticality' . . . . . . . . . . . . . . . . . . . 102
5.3 The simple sentence in English . . . . . . . . . . 104
5.4 Categories . . . . . . . . . . . . . . . . . . . . . . . 107
5.5 Constituency . . . . . . . . . . . . . . . . . . . . . 112
5.6 Structural Ambiguity . . . . . . . . . . . . . . . . 116
5.7 Some Problems . . . . . . . . . . . . . . . . . . . . 118
5.8 D-structure, S-structure, and Transformations . 121
5.9 Yes-No Questions . . . . . . . . . . . . . . . . . . 124
5.10 Summary . . . . . . . . . . . . . . . . . . . . . . . 131
5.11 Exercises . . . . . . . . . . . . . . . . . . . . . . . 131


## 5.1 Introductory remarks

Having surveyed the structure of sounds (phonetics & phonology) and of complex words (morphology), we now turn to the study of the structure of sentences. Actually, as far as our every day experience of language goes, we are generally confronted not by individual sounds, nor by individual words, but rather by strings of words, put together in a manner made possible by the syntactic system of the grammar in question. It is therefore the sentence which represents the most salient aspect of language in actual use.

## 5.2 'Grammaticality'

It is apparent that some strings of words represent possible sentences for a given grammar, while other strings of words do not. Compare the strings in (5.1a) and (5.1b) below, for example.

(5.1) a. The clown likes only fresh twinkies.
b. *Twinkies only the likes fresh clown.

It is apparent to anyone with a grammar which permits them to read this book that while the string of words in (5.1a) represents a *sentence* of English, those same words arranged as in (5.1b) do not. In particular, (5.1a) appears to have a relatively straightforward interpretation (it asserts of some specific clown that he has the property of liking only fresh twinkies, as opposed to stale twinkies or fresh asparagus).[1] By contrast, the string of words in (5.1b) does not appear to assert anything about anyone (or anything). One can easily imagine contradicting (5.1a): 'no, no, he likes stale twinkies as well,' for example, but what could one say to (5.1b) except 'what are you trying to say?'

It is traditional in modern syntactic literature to refer to a string of words which speakers with the relevant grammar consider to be a 'sentence' as a *grammatical* string or a grammatical sentence, while a string of words which does not have this property is generally called *ungrammatical* and is marked with an asterisk, as in (5.1b) above.

It is important to try to understand precisely what the term 'grammatical' means in modern linguistics – as with other terms we have seen, the intended force of this term is technical and highly specific. First, grammaticality is not a question of whether or not you have ever heard a particular sentence or whether the string involved is one you are likely to say. For example, it is probable that you have never said the sentence in (5.2).

[1] Note that the clown in question may or may not like fresh asparagus, the sentence says nothing about the matter one way or another.

(5.2) I just whacked my little toe against the P-Z volume of the Oxford English Dictionary again!

It may be for many of you that it is extremely unlikely that you have ever heard or would ever say such a sentence – e.g., you may not have the Oxford English Dictionary, or you may be really careful when you walk around barefoot. Nevertheless, even for those of you for whom utterance of such a string is virtually unthinkable, it is clearly a perfectly well-formed sentence for speakers with English-type grammars. (Syntactic well-formedness is synonymous with grammaticality.) Well-formedness, therefore, has nothing to do with likelihood of utterance. It is also the case that well-formedness has nothing to do with length of utterance – a well-formed sentence may be arbitrarily long. For example, if we had you make up the longest sentence you could think of, we could simply add 'We know that' to the beginning of that sentence, and the result would be well-formed. You, in turn, could add 'You know that' to the beginning of our sentence (so that we get 'You know that we know that...') and we could all go on making the sentence longer and longer in this manner for days, weeks, or years on end. The resulting sentence would be boring to listen to and kind of useless, but it would be *well-formed*, and therefore grammatical. The likelihood that anyone is going to say a sentence that it would take us years of dedicated effort to construct is, we are happy to say, vanishingly small.

Secondly – and this is sometimes a difficult point for non-linguists to fully appreciate – grammaticality is about *syntactic* well-formedness *only*. It therefore does not matter if the string of words makes sense at all when one is attempting to determine whether that string is grammatical. Look at the examples in (5.3).

(5.3) a. The refrigerator hid under the mouse.
b. Excitement said that the television had a new hairdo.
c. My brother is pregnant again.

Contrast these with the parallel sentences in (5.4).

(5.4) a. *Under mouse the hid refrigerator the.
b. *Said the had new a that excitement television hairdo.

c. *Pregnant my is brother again.

Clearly, the examples in (5.3), while semantically odd, are well-formed in a way in which those in (5.4) – which involve exactly the same words – are not. One can easily imagine somewhat fantastic stories in which each of the sentences in (5.3) might occur (like 'the tale of the tiny animate refrigerator,'), but the strings of words in (5.4) cannot be part of any story, no matter how fantastic.[2] Notice that the oddness of the sentences in (5.3) is essentially oddness about the world – refrigerators don't normally 'hide,' excitement can't 'speak,' TVs don't have hairdos, and men don't get pregnant. The world, and human conceptions of it, change all the time. It would be quite a bad thing, therefore, if your grammar did not permit you to say things just because they were not *currently* sensible. Language is flexible enough to serve your needs whatever bizarre situation you happen to find yourself in.

Basically, for at least certain types of words (nouns, for example), syntactic well-formedness has nothing to do with the 'meaning' of the word in question. Your grammar does not care whether you are talking about 'dogs' or 'refrigerators' – you can construct exactly the same sentences using either noun.[3] The same may be said for adjectives, e.g. – every sentence about 'a brown refrigerator' is equally well-formed if instead we say 'a large refrigerator' or 'a skeptical refrigerator' (though those formed with the latter phrase may be semantically odd). This is the contrast between syntax and semantics: syntax is about syntactic well-formedness, semantics is about meaning.

## 5.3 The simple sentence in English

The basic unit of syntactic analysis is the sentence. A (grammatical) sentence is a string of words which the syntactic component of the grammar can assign a coherent structural analysis to. The strings in (5.1a), (5.2), and (5.3a-c) all represent sentences, while those in (5.1b) and (5.4a-c) do not. A sentence is made up of words, which, as we saw in our discussion of morphology, may be either (morphologically) simple or complex. As the constrast between (5.3)

[2]Unless, of course, the fantasy is that the language of the story in question is no longer a variety of English. But you would be pushing it . . .

[3]A noun (just in case) is a word which designates an entity, perhaps abstract. Given a sufficiently liberal notion of 'thing,' the traditional definition ('a noun is a word which designates a person, place, or thing') is adequate for our purposes.

and (5.4) suggests, in English-type grammars at least, the arrangement of words is critical as to whether or not a given string represents a (grammatical) sentence or an (ungrammatical) string.[4]

As we stated, the semantics of the individual words does not play a key role in determining syntactic well-formedness. What *does* turn out to be extremely relevant to this determination is the so-called *syntactic category* of the word in question. While we can substitute virtually any noun for 'dog' in the sentence 'the dog is ill-behaved' – giving 'the cat is ill-behaved,' 'the tree is ill-behaved,' 'the nation is ill-behaved,' 'the excitement is ill-behaved,' etc. – if we attempt to place a verb (like 'drowns' or 'believes') in the position of the noun 'dog', we get ungrammatical strings: '*the drowns is ill-behaved' and '*the believes is ill-behaved'. (Recall that ungrammatical strings are marked with an asterisk).

Also clearly relevant to grammaticality are certain morphological properties of the words in question. Examine the sentences and ungrammatical strings in (5.5).

(5.5) a. The dog sits on the carpet.
b. *The dog sit on the carpet.
c. The dogs sit on the carpet.
d. *The dogs sits on the carpet.

In general, verbs must agree with the *number* of their subject in English-type grammars (though there is some variation in the dialects).[5] That is, if the subject is singular ('the dog') the verb must be singular ('sits'), while if the subject is plural ('the dogs') the verb must be plural ('sit'). Such agreement processes are common in human languages. Agreement will not be a major concern of ours in our brief survey of syntactic processes – but it is quite important in advanced syntactic theorizing.

What does one need in order to have a sentence? Examine the strings in (5.6).

[4]The latter is sometimes called an ungrammatical 'sentence' which is something of a contradiction in terms, given what we have just said.

[5]The 'subject' is what the sentence makes an assertion about – in the sentences in (5) it is one or more dogs. We will return to this issue shortly.

(5.6) a. *The dog on the carpet.
b. *Sits.
c. *The sits.
d. The dog sits.
e. The dog sits on the carpet.
f. Dogs sit.
g. Dogs sit on carpets.

Why are the strings in (5.6a-c) ungrammatical? Superficially, the contrast between (5.6a) and (5.6e) would seem to indicate that the problem with (5.6a) is that it lacks a *verb*. Similarly, the contrast between (5.6b) and (5.6f) would seem to indicate that the problem with (5.6b) is that it lacks a subject. (5.6c) tells us that 'the' is not a possible subject (and thus (5.6c) doesn't fix the problem with (5.6b)) – whereas (5.6d) tells us 'the dog' *is* a possible subject. It appears, then, that the minimal sentence will have an appropriate subject and a verb.[6] This makes sense, of course. Sentences basically predicate, or attribute, some property (expressed by a verb and its satellites) of some entity (expressed by a noun or pronoun and its satellites). Thus (5.6f) predicates of the class of entities designated as 'dogs' that they have the property of engaging in sitting. (5.6g) asserts that entities known as 'dogs' have the property of sitting on entities known as carpets. (5.6d), by contrast, says only that some specific dog – known to the speaker and hearer (this is what 'the' does) – has the property of sitting. Note that it does not say that dogs in general *don't* sit, it is silent on the general properties of dogs. Every declarative[7] sentence predicates some property of some entity.

(5.6c) indicates that there are restrictions on what can be a subject. Typical types of subject can be seen in (5.7).

(5.7) a. The dog sits on the carpet.
b. The big dog sits on the carpet.
c. The big dog in the corner sits on the carpet.

[6]We will consider a little later *imperative* sentences such as 'Leave!' which appear to lack a subject.

[7]We will consider other types of sentences – interrogatives and imperatives, for example, in what follows.

d. She sits on the carpet.

What is interesting about these examples is that in (5.7a-c) there is always a *noun* present in the string that represents the subject ('the **dog**'; 'the big **dog**'; 'the big **dog** in the corner') and that if we leave that noun out, the strings become ungrammatical (*the sits on the carpet; *the big sits on the carpet; *the big in the corner sits on the carpet). In (5.7d) there is a pronoun (a substitute for a noun phrase), which is also (in English-type grammars) obligatory in its sentence.

Similarly, acceptable predicates appear to require a verb.

(5.8) a. The dog **sits**.

b. The dog **is** fat.

c. The dog **drinks** the water.

Again, without their verbs, these strings are not sentences (*the dog; *the dog fat; *the dog the water).[8]

It appears, then, that a sentence consists of two basic parts: a subject, which designates some entity in the universe about which something is going to be said, and a predicate, which designates some property which the sentence asserts is true of the entity designated by the subject. The subject must contain a noun or a noun substitute (a pronoun); the predicate must contain a verb.

These are not properties of English sentences, only. As far as we know, these are universal properties of natural language. One might say, instead, that to qualify for 'sentencehood' in a human brain, a string must contain both a subject and a predicate.

## 5.4 Categories

As we have seen, there are apparently restrictions on what can function as a subject of a sentence and on what can function as a predicate. It makes sense,

[8] Note that it does not follow that you cannot *say* the strings. If someone asks 'Who sat on the carpet?', the answer 'the dog' is perfectly acceptable. You might, however, consider what the answer 'the dog' means – e.g., can it mean 'the dog runs' or 'the dog sleeps on the carpet'?

if we are to try to understand how sentences are constructed, to explore these restrictions – what makes for an acceptable sentence is undoubtedly in part a function of what makes for acceptable components (subject and predicate) of a sentence.

We will first take a look at acceptable subjects. In (5.6f) the subject is simply the noun 'dogs.' We can conclude from this that a noun alone can be a subject. We will designate nouns in general by the symbol N. In (5.7a) the subject is 'the dog.' 'the' is what linguists call a 'determiner' (like, in English, 'a(n),' 'his', 'my', 'that,' 'these,' etc.). We will designate determiners by the symbol D. An acceptable subject thus may be either N or D N. One of the ways to think about the relationship between N and D N is that a determiner *may* be present ('the dogs') but is not required ('dogs') – i.e., a determiner is *optional*. Notice that the noun is not optional (*the is happy, *my is here). If we designate optional material by placing it in parentheses, we can say that a subject may have the form '(D) N' – that is, an optional determiner followed by a noun. The subject of (5.7b), 'the big dog,' contains, in addition to a determiner and noun, an adjective, 'big.' If we designate adjectives with the symbol A, we can now say that a subject may have the form '(D) (A) N' – i.e., an optional determiner, an optional adjective, and a noun.

It is not entirely accurate, however, to say that '(D) (A) N' is a generalization about *subjects*. Examine the sentences below.

(5.9) a. The cat sits on dogs.
b. The cat sits on the dogs.
c. The cat sits on the big dogs.

These examples indicate that the same N, (D)N, or (D)(A)N pattern occurs in non-subject positions. In (5.9), they are the object of the preposition[9] 'on'. Given that this, too, is possible, we do not want to connect N, (D)N, and (D)(A)N only to subjects. It turns out, in fact, that these same forms occur in many different positions in the sentence. Interestingly, wherever you can have an N, you can also have the other forms which include an N. We can make a generalization about this by calling an N and its satellites a

[9]A preposition is a word which expresses spatial, temporal or causal relationships – typical examples from English include *in, on, over, above, under, through, down, across,* etc.

Noun Phrase or NP, for short. It thus appears that Noun Phrases (NPs) are fundamental building blocks of the sentence. They are required in subject position, but they also occur elsewhere (as 5.9a-c show).

We have not yet considered the subject of (5.7c), 'the big dog in the corner.' Notice that although this subject contains *two* nouns, only one of them is the real or structural subject (namely, the 'dog' – i.e., it is the dog in (5.7c) which is sitting on the carpet, not the corner that is doing the sitting). 'the corner', by contrast, is the object of the preposition 'in'. Let us designate prepositions with the symbol P. Notice that '*the dog in sits on the carpet' is ungrammatical. This is because 'in' requires an object when used as a preposition (prepositions are pre-posed to their nominal objects). We can therefore posit another type of a phrase, call it a Prepositional Phrase (PP) which consists of a P and an NP (i.e., a preposition and its noun phrase object). A Noun Phrase, as (5.7c) shows, can also contain a PP – thus the structure of Noun Phrases appears to be the following: (D) (A) N (PP). One way to express the 'X consists of Y' is by the use of what are called Phrase Structure Rules. For example, the following Phrase Structure Rules express the observations we have made thus far about NPs and PPs:

(5.10) a. NP → (D) (A) N (PP)
b. PP → P NP

The interpretation of these Phrase Structure Rules is as follows: (5.10a) says that a Noun Phrase consists of an optional determiner, an optional adjective, a noun (not optional) and an optional Prepositional Phrase. (5.10b) says that a Prepositional Phrase consists of a Preposition (not optional) and a Noun Phrase (also not optional).

Note that just with these two rules we can begin to see how we might account for some of the remarkable properties of syntactic systems – how they allow us to generate such a wide variety of sentences (given the limited storage space we have in our brains) and how they allow us to generate infinitely long strings (given the same problem). With just these two very modest rules, conjoined to the relatively small lexicon in (5.11), we can produce all the well-formed phrases in (5.12), and infinitely many more.

(5.11) D: the, my, that, his

A: stale, large, yellow, disgusting

N: twinkie, clown, man, dog, twinkies, dogs

P: in, on, by, of

(5.12) a. the stale twinkie on the large clown

b. twinkies

c. the disgusting dog by the stale man

d. the dog by the clown on the stale twinkie of the disgusting man

All of the phrases in (5.12) make perfectly good subjects in English-type sentences (as well as perfectly good objects and objects of prepositions). In addition, it is obvious that the string in (5.12d) could be extended to infinite length by simply continuing to add Prepositional Phrases to the end of it. This is because the rules in (5.10) have an interesting structural property: they are *recursive*. (5.10a) is a rule about how to form NPs and it contains the provision that an NP may contain a PP. The rule in (5.10b), in turn, is a rule about PPs, but notes that all PPs must have within them an NP. Thus we can use these rules to generate an NP which contains a PP (since it can, by 5.10a) which contains an NP (by 5.10b) which contains a PP (by 5.10a) which contains an NP (by 5.10b) which contains a PP (by 5.10a) which contains an NP (by 5.10b) – no doubt you get the picture ... This is *recursion* – it provides the grammar with a means to produce infinitely long strings from very finite means.

We have not yet discussed the properties of predicates, aside from noting that they must contain a verb and that they are a necessary part of any grammatical sentence. Since predicates *must* contain a verb and *may* contain additional elements, they are, logically enough, Verb Phrases (i.e., VPs). We can capture the fact that they are required in well-formed sentences by positing a Phrase Structure Rule which says that a Sentence (which we can abbreviate S) consists of a NP (its 'subject') and a VP (its 'predicate') – i.e., S $\rightarrow$ NP VP. What can a simple VP consist of? Some examples are given in (5.13).

(5.13) a. walks

b. eats the twinkie

c. goes to the party

d. throws the twinkie to the clown

In the first example we see the minimal VP – a verb and nothing more. In (5.13b) we see a verb followed by a so-called 'direct object,' which has the form of an NP. Note that any NP can go here ('eats the stale twinkie,' 'eats the large dog,' etc.). In (5.13c) we have a verb which is followed by a PP – again, any PP will do here ('goes in the lake,' 'goes by the house,' etc.). Finally, in (5.13d) we have a verb which is followed by both an NP ('the twinkie') and a PP ('to the clown'). Everything but the verb is optional, as (5.13a) shows, so we can write the Phrase Structure Rule for the VP in English as in (5.14d). (We will repeat the Phrase Structure Rules mentioned earlier here for convenience.)

(5.14) a. S → NP VP

b. NP → (D) (A) N (PP)

c. PP → P NP

d. VP → V (NP) (PP)

What these Phrase Structure Rules (PSR's) of English say is this: a sentence (S) consists of a Noun Phrase (NP) and a Verb Phrase (VP), both of which are obligatory. A NP consists of an optional determiner (D), an optional adjective (A), a noun (N) – which is obligatory – and an optional Prepositional Phrase (PP) in that order. A PP consists of a preposition (P) and a NP (both obligatory). A VP consists of an obligatory Verb (V), an optional NP and an optional PP.

There is an additional convention for PSR's which is convenient to know and that is how to indicate multiple instances of an element. For example, we are not limited to a single adjective in our description of nouns. Actually, there appears to be no *structural* limit on number of adjectives. As with recursion and other mechanisms which allow production of infinitely long strings, the limit appears to be imposed by extralinguistic factors, e.g., the patience of the listener, short-term memory, etc. We indicate that iterations of a category are possible by placing an asterisk after the category designation. For adjectives, this is illustrated below.

(5.15) NP → (D) (A)* N (PP)

Now our NP rule will produce such strings as:

(5.16) the old, stale, disgusting twinkie

Not all categories can be repeated in this manner. You should test to see which others can and cannot be repeated.

While the rules above do not exhaust the possible syntactic structures of English, they do account for a very large percentage of such structures. We will not attempt to account for the remainder here. It is important to note, however, that we are using PSR's to model speakers' knowledge of syntactic structure. Since we have been assuming for every aspect of the grammar that speakers are *competent* (in spite of their frequent performance errors), we want our PSR's to produce *all and only the grammatical strings*. Each time we revise our PSR's to take some new data into account, we need to check to make sure that we have not altered them so that they *overgenerate* – that is, produce ungrammatical as well as grammatical strings.

## 5.5 Constituency

Speakers' intuitions about what strings can be sentences and what strings cannot be sentences led us to posit two, main *constituents* for every sentence, a NP and a VP (corresponding to a subject and a predicate).

(5.17) The dog with bulging eyes chased the two bold cats.

The internal structure of sentences turns out to be more complex than that, however. There are sub-constituents embedded within each main constituent, and, depending upon the complexity of the sentence, there are further sub-constituents. What leads us to believe this? Again, it is speakers' intuitions which suggest that sentence have this type of internal structure. Since speakers' intuitions appear to arise from their knowledge of language, we want our model of syntax to account for those intuitions.

There are a number of tests for categories and constituency – one of these is conjunction.

(5.18) The dog [sits on the sofa] and [sleeps on the bed].

(5.19) The dog sits [on the sofa] and [on the bed].

(5.20) The dog sits on [the sofa] and [the bed].

In each of these examples, two elements belonging to the same syntactic category have been conjoined – two VP's, two PP's, or two NP's. Interestingly, it seems that only like categories can be conjoined by 'and.'

(5.21) *The dog bites with his teeth and the mailman.

This is further support for the hypothesis that the syntactic component is sensitive to categories such as N, PP, and so forth. In addition, it suggests that groupings of certain elements, such as PP's, form sub-constituents within the sentence.

Substitution tests also suggest that the sentences consist of structures embedded within other structures – that is that they are **hierarchical** in nature.

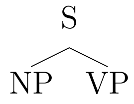
(5.22) The dog sits on the carpet and the cat sits there, too. [there = 'on the carpet']

(5.23) The dog sits on the carpet and the cat does, too. [does = 'sits on the carpet']

In the above, 'there' substitutes for a PP and 'does' substitutes for a VP. This indicates that the PP 'on the carpet' forms a syntactic grouping, a constituent, *within* the VP. Similarly, the VP forms a syntactic grouping, or constituent, within the S (as we have already determined).

It is probably easiest to see the relationships between constituents of sentences if we display them graphically using syntactic **trees**. Trees are typically used to represent syntactic relationships within particular sentences. Note, though, that the structures for every sentence must be able to be generated by our PSR's which means that every tree must be *consistent* with our PSR's. We can 'translate' the information from the PSR's into tree structures in the following way. The material to the left of the arrow in a PSR is a dominating node while the material to the right of the arrow represents the nodes dominated by that dominating node. For example, the tree which corresponds to our first PSR is given below.

(5.24)

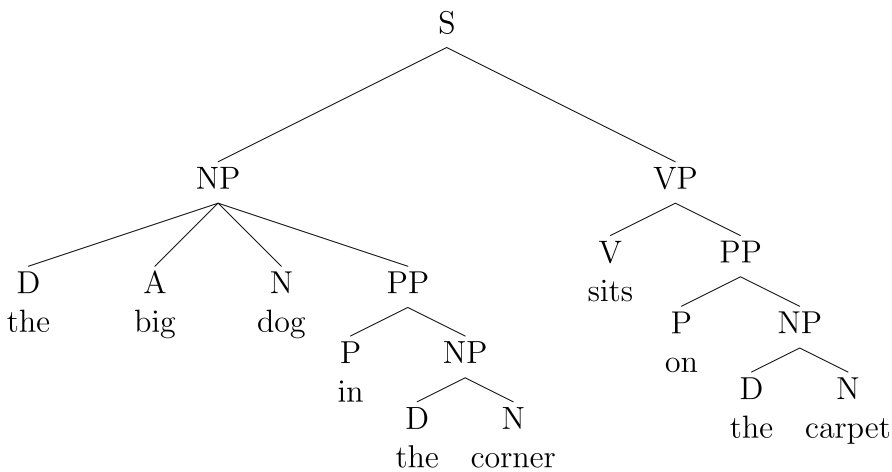



In the above, S is said to *dominate* both NP and VP. The traditional terminology used of trees is that of female genealogy: a node which dominates other nodes is a *mother* of those nodes, the nodes are *daughters* of their mother and they are *sisters* of one another. The NP and VP in this example, therefore, are sisters. We stated earlier that the requirement for 'sentencehood' was the presence of an NP (subject) and a VP (predicate) – not just for English but for human language. Since this is the case, every sentence of every natural language will have these two main constituents. The only (superficial) differences among languages will be in linear order. Only two orders are possible, NP VP and VP NP. Both orders are attested in natural language.

Beyond this, we can give an example of the tree for an actual sentence (one that we saw earlier in 5.7c). Any particular sentence will be unlikely to use all of the options allowed by the PSR's, especially if we had a more complete set of PSR's. For instance, the following sentence only has one adjective and does not have a NP direct object. Again, it is *consistent* with our PSR's – all of the obligatory elements are present and all of the relationships are legitimate – it simply uses some, but not all, of the options available.

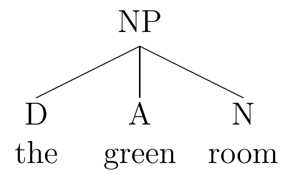
For example, the tree for the sentence in (5.7c) is given below.

(5.25)

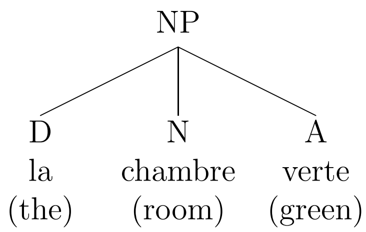


Each *phrasal category*, that is, each element whose name has the form XP (NP, VP, PP, etc.) has a **head** – the element which is obligatory. All NPs must contain a noun (or a noun-substitute, a pronoun), all VPs must contain a verb. Thus the N is the head of the NP, the V the head of the VP, the P the head of the PP, and so on. All of the non-head elements within the phrasal category *modify*, or supply more information about, the head of the phrase. The structural relationship between the non-head elements, or modifiers, and the head is one of **sisterhood**. Thus, in the leftmost NP in the tree, *dog* is the head. Its sisters (and therefore modifiers) are the D *the*, the A *big*, and the PP *in the corner*. Note that *in* is NOT a modifier of *dog* because it is not a sister of *dog*. It follows from the fact that *big* is a modifier of *dog* that this sentence is about a big dog (not about a big carpet or a big corner).

Our common intuition that the words of a sentence fit together in a particular way – that, e.g., 'big' is a modifier of 'dog' rather than of 'carpet' in (5.17) – is therefore captured *structurally* in the tree in (5.17). The relationships illustrated by both trees and PSR's have two components. The first is that of hierarchical structure – dominance and sisterhood among constituents of a sentence. The second is *linear* order. The words of every spoken sentence are uttered over time and not simultaneously. This constraint alone forces a linear order.[10] Languages differ with respect to the linear order of constituents but *not* with respect to hierarchical structure – all natural language syntactic structure is hierarchical in nature. Some languages postpose adjectives, for example, instead of preposing them as in English. A direct translation of the phrase 'the green room' from the French will result in 'the room green' because (most) adjectives are postposed to their nouns in French. Examine the trees for a simple NP in English and French:




[10]It could well be that the syntax is not constrained in this way but the fact remains that there is a bodily constraint of this type.




Preposed or postposed, the relationship between the adjective and the noun (the adjective is sister to the noun and the noun is the head of the NP) is the same.

## 5.6 Structural Ambiguity

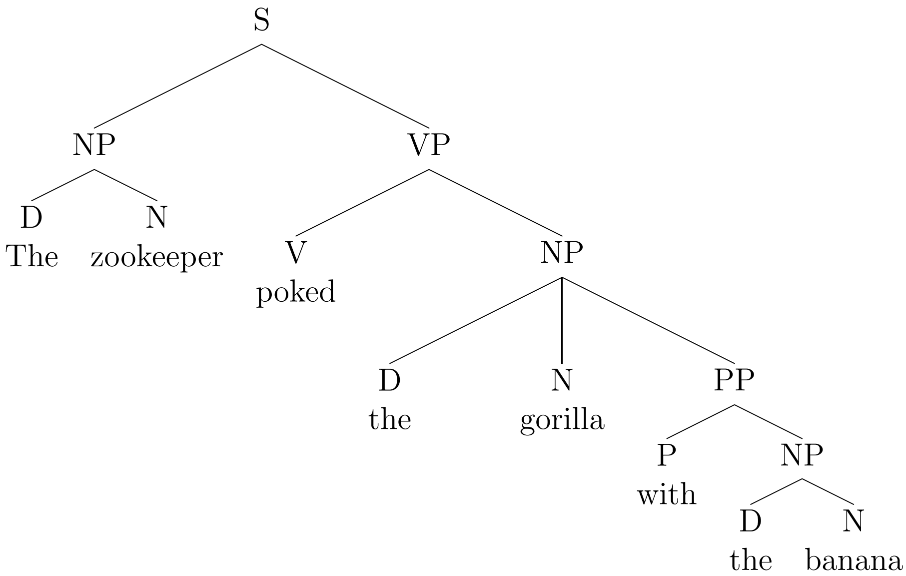
We saw in Chapter 4 that morphologically complex words may be structurally ambiguous – have more than one interpretation due to more than one possible structure. (It was noted at that time that the ambiguity exists only from the listener's perspective – the speaker *knows* which structure he or she generated.) Sentences can be structurally ambiguous in virtually the same way except that the constituents involved are syntactic categories, not complex morphemes. Consider the following sentence in (5.26).

(5.26) The zookeeper poked the gorilla with the banana.

This sentence is ambiguous – it has two possible interpretations. One of the interpretations is that the zookeeper is using a banana as an instrument to poke the gorilla. The other interpretation is that the gorilla is holding a banana and the zookeeper is poking the gorilla (but we do not know with what – his finger, a different banana, whatever . . . ). Critically, you only had to read or hear the sentence once (and not twice) to get these two interpretations. Since this is true, we know that nothing about the sentence changed to allow a different interpretation from one point in time to the next. How do we account for this? The answer lies in the different possible syntactic structures that a listener can assign to this sentence when they hear it.

Syntactic trees express relationships between the constituents of a sentence. Those elements which are most closely related are joined as sisters. If we think about the ambiguity of (5.26), we can see that it revolves solely around 'with the banana', more specifically, around whether or not the gorilla is holding the banana. If we want to indicate that the gorilla is holding a banana, we must express the close relationship between those two constituents – the PP 'with the banana' and the N it is modifying, 'gorilla'. Such a relationship can be expressed with the following structure.

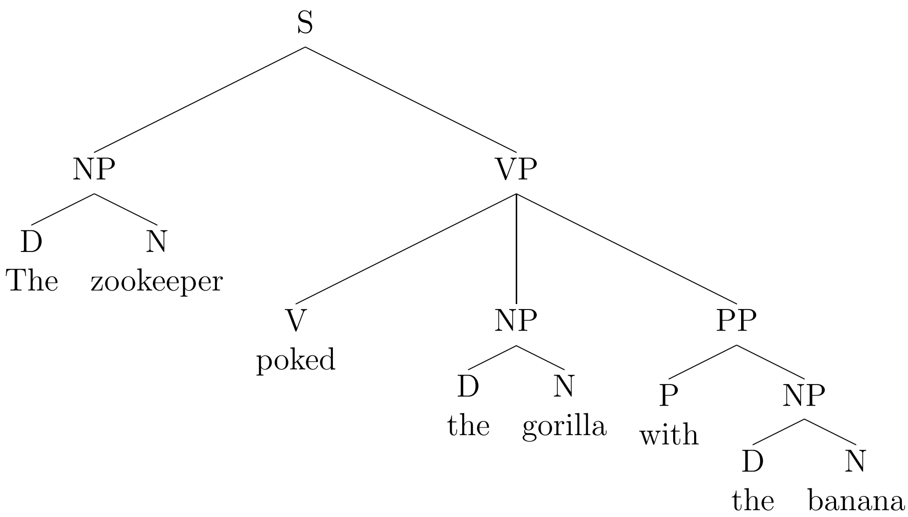
(5.27)




This tree indicates that the PP 'with the banana' modifies the head N, 'gorilla' – the PP is sister to the N. We are talking about a specific gorilla, here, one with a banana.

On the other hand, if the gorilla is not holding the banana, we do *not* want to express a close relationship between the PP and the N – the PP is no longer modifying the N. In this case, the tree would be that in (5.28) below.

(5.28)




Note that in the above tree, the PP is sister to the V. This means that the PP is modifying the V. This is the result that we want. For the interpretation where the gorilla is *not* holding the banana (5.28), what we know for sure is that the poking activity of the verb is being carried out by means of a banana as instrument – that is, 'with the banana' is modifying 'poked'. Since the zookeeper, as subject, is doing the poking and the poking is being carried out by means of a banana, it is likely, but not definite, that the zookeeper is holding a banana.

Cases of structural ambiguity illustrate the importance of understanding that trees and phrase structure rules indicate *relationships* between the elements of a sentence.

## 5.7 Some Problems

Obviously, although the PRS's we've posited for English-type grammars above account for a large number of sentences, there are also a very large number that are not accounted for. Some of these require only slight revisions to the PSR's above,[11] which we can safely ignore in an introductory

[11]For example, the fact that you can say 'the very big dog' instead of 'the big dog' indicates, given that 'very' modifies 'big' rather than 'dog', that instead of 'A' in the NP phrase structure rule above we should probably posit an adjective phrase (i.e., AP). The AP itself would consist of an optional adverb (Adv) and an obligatory A (the head).

text. Others appear to be more problematic. For example, look at the sentences in (5.29).

(5.29) a. Be happy!
b. Sit on the carpet!

These examples seem to violate our first and, in many ways, our most fundamental Phrase Structure Rule: S → NP VP. In particular, they appear to consist *only* of a VP. We could revise our PSR's so that our initial rule said: S → (NP) VP, but there are reasons for thinking this is not a good idea. First, note that if a VP alone can be a sentence, it is a little mysterous why strings like **goes to the store* or **went to Montréal* are not grammatical. A second consideration can be seen in the examples in (5.30).

(5.30) a. I kick myself.
b. *I kick yourself.
c. *I kick himself.
d. *Kick myself.
e. *You kick myself.
f. You kick yourself.
g. *You kick himself.
h. Kick yourself.

Notice the patterning of these sentences, all of which involve a *reflexive* pronoun (-self). (5.30a) is fine, but (5.30b) and (5.30c) are bad. Why? Looking at (5.30efg) is helpful – (5.30f) is good, but (5.30e) and (5.30g) are ungrammatical. It appears that in a sentence with a reflexive pronoun object, the subject and object must refer to the same person. This is strongly supported by the pair of sentences in (5.31).

(5.31) a. He$_i$ kicks himself$_i$.
b. *He$_i$ kicks himself$_j$.

These sentences employ a convention used in the study of syntax by which the referent of a pronoun (or NP) is indicated by a *subindex*, usually italic *i, j, k*... The indices in (5.31a) indicate that the object in the world referred to by the subject pronoun 'he' is *i* (someone called John, for example) and that the object in the world referred to by the object pronoun 'himself' is also *i* (also, therefore, John). The two pronouns are thus *correferential*. By contrast, in the ungrammatical sentence (5.31b), the subject 'he' refers to one individual in the world *i* (John) and the object 'himself' refers to another *j* (whom we will call Bill). The reflexive pronoun is only grammatical in the first case. (Just for fun, what happens when we keep all the indices the same in (5.31ab), but we replace 'himself' with 'him'?)

It is now apparent that if you have a reflexive pronoun direct object, it must be correferential with the subject of its verb. (5.30b) is ungrammatical precisely because 'I' cannot be the same object in the world as 'you.'[12] But then why is (5.30d) ungrammatical, but (5.30h) grammatical? It seems like 'Kick yourself!' acts as if it had a subject pronoun 'you,' even though it doesn't *seem* to.

This corresponds, no doubt, to what you were probably told in school about such sentences: that they have an 'understood' subject that is *you*. It is unlikely, however, that your teacher told you what it means for 'you' to be 'understood' as the subject of such sentences. One way to make this notion coherent is to say that the sentence is *interpreted* as having a 'you' instead of having nothing as its subject NP. Note that this is not a general property of English – for example, the 'missing' object of (5.32b), as contrasted with (5.32a), cannot be interpreted as 'you.'

(5.32) a. I kick Newt.

b. I kick.

Sentence (5.32b) cannot mean 'I kick you,' in spite of the 'missing' object. So it is not the case that any missing NP can simply be interpreted as 'you' in English. In fact, as the sentence 'be happy' tells us (note **are happy*), subject 'you' can only be 'left out' in the pronunciation of a sentence if the

[12]'Object in the world' is to be interpreted in the usual way – i.e., the world in question is the mental world, not the actual physical one. This allows one to use reflexive pronouns in sentences such as 'the unicorn admires itself in the mirror,' where there is no real-world object involved.

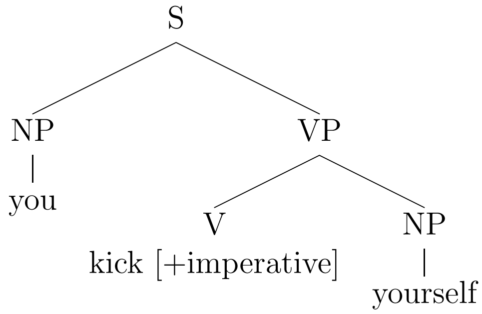
verb in the sentence is an *imperative* (command) verb. *you went to the store* is grammatical, but the string **went to the store* cannot be interpreted with an 'understood you' subject, because 'went' is not an imperative verb. These observations will be helpful in solving the mystery of these VP-like sentences, to which we now turn.

## 5.8 D-structure, S-structure, and Transformations

English-type grammars appear to have the following property: if you have an imperative verb in a sentence (let us assume that imperative verbs all have a feature, like [+imperative]) and there is no subject, interpret the sentence as if there were a subject 'you.' Basically, this means that the listener supplies a 'you' in just this circumstance. But what about the speaker? How does the speaker ensure that his or her grammar produces *go to the store* for *you go to the store* but not **went to the store* for *you went to the store*? Since the grammar is designed to capture simultaneously what you do as a speaker and what you do a listener, we must unify what we have now decided the listener does (add a subject *you* if the verb is imperative) with what a speaker must do. To take a parallel from phonology, if a speaker intends to say the word /bin/ 'bean' their grammar will add nasalization to the [i] (because of their phonological rules), producing [bĩn]. If they hear someone say [bĩn], on the other hand, they will reverse this process, and instead of adding nasalization (which is what they do as a speaker), they will take it away. (The same unconscious knowledge of the grammar provides for a necessary two-way street in this regard.) Thus, in some sense, speaking is the opposite of listening: as a listener one subtracts the nasalization which, as a speaker, one would add.

In the case of imperative subject 'you' then, since we know the listener *adds* a 'you' to the sentence (i.e., 'understands' *you* to be the subject, even though it was not uttered), as a speaker it would make sense if we started with 'you' in the sentence, and *deleted* it (thus doing the opposite of what the listener does). The speaker's original representation of *kick yourself* will thus be that given in the tree in (5.33).

(5.33)



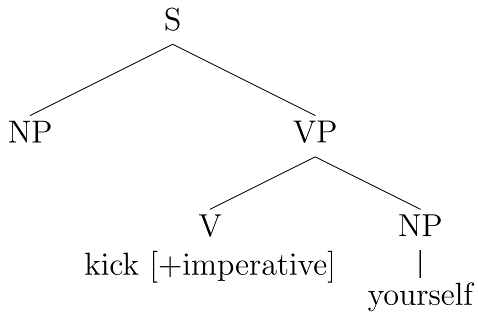

What the grammar does to this sentence can be expressed by the following rule, which we will call the Imperative Transformation.

(5.34) **Imperative Transformation:**

- delete 'you' if it is the subject of an imperative verb

The result of this transformation is the sentence in (5.35).

(5.35)




Since the listener has this same knowledge (notice that someone who does not have an English-type grammar cannot tell that a 'you' is to be understood in a sentence like 'kick yourself'), he or she knows that when they hear a sentence with an imperative verb and no subject, the subject must have been deleted by the Imperative Transformation. The listener can thus restore the tree in (5.35) to its original status in (5.33) by reversing the Imperative Transformation process.

This means that in the course of generating the sentence 'kick yourself', the grammar has two different representations of the string. The original representation (with the ‘you,’ as in 5.33) matches the ‘interpretation’ of the sentence in some meaningful sense, before things that will be ‘understood’ by the listener have been deleted. The final representation, which will be sent to the phonological component of the grammar to be pronounced, may have undergone the effects of rules such as the Imperative Transformation. The structure before any transformations have applied to it is known as the Deep Structure, or, nowadays, as the d-structure. The string that results from the applications of the various transformations (we have only seen one so far, but we’re about to see more) is known as the Surface Structure, or, nowadays, as the s-structure. The mapping from d-structure to s-structure is regulated by the transformational rules found within the grammar. The structure of the syntactic component of the grammar is therefore as given in (5.36).

(5.36)

$$\begin{tikzcd}
\boxed{\text{lexicon}} \arrow[d, "\text{PSR's}"] \\
\boxed{\text{d-structure}} \arrow[d, "\text{transformations}"] \\
\boxed{\text{s-structure}}
\end{tikzcd}$$

This diagram should be interpreted as a model of speaker production. (For comprehension, reverse the arrows and add ‘undo’ to ‘transformations.’) Speakers have lexicons (mental dictionaries) from which they extract elements for a sentence. Those lexical items are organized into a syntactic representation which is in accord with our PSR’s. This underlying representation, d-structure, is fed into the syntactic portion of the computational system. Any transformational rules which apply to the d-structure are applied, producing the surface structure (s-structure). That’s the end of the syntactic part of the operation. The similarities between this and the phonological and morphological components should now be very clear. For all three components of the grammar, there are two levels of representation, one is more abstract (the underlying representation) and the other, less abstract, is the output of the relevant portion of the computational system.

## 5.9 Yes-No Questions

The key features of transformations, as can be seen from the example we have given, is that they do not affect the interpretation of the sentence to which they apply (since the listener 'undoes' their effects) – they apply for purely structural reasons.[13] We can be at least a little more precise about what we mean by 'structural reasons,' using imperatives as an example.

Following our reasoning above, the d-structure for an imperative contains a pronominal subject 'you' (which is subsequently deleted by imperative transformation). We now have the possibility of identical pairs of strings at d-structure, only one of which results in an imperative. A potential example is given below.

| | | | |
|---|---|---|---|
| | d-structure | you kick the gopher | you kick the gopher |
| | | $\Downarrow$ | $\Downarrow$ |
| (5.37) | transformations | Imperative Tr. | No transformations |
| | | $\Downarrow$ | $\Downarrow$ |
| | s-structure | Kick the gopher! | You kick the gopher. |

A critical question arises here. What has caused one of these strings to come out as an imperative and the other to come out as a declarative (i.e., statement)? Because we believe that the computational system is entirely systematic in its treatment of input – that is, given identical inputs, it will produce identical outputs – we are forced to posit at least one additional element for at least one of the above strings. Minimally, we could say that when the intention is to generate an imperative, a special 'imperative marker' (which has no phonological content – something like a 'zero morpheme') is drawn from the lexicon along with the other lexical items. It is this 'imperative marker' which triggers the imperative transformation. One possibility is that markers of this type must be included in the construction of every sentence – markers which indicate the 'force' of the utterance, declarative, imperative, interrogative, and so on.[14] We conclude, then that transformations must be motivated by the presence of a trigger in the string.

What other sorts of transformations are there? Well, keeping in mind that we were discussing the force of utterances, consider the examples below.

[13] This is the same as phonology: the nasalization added to /bin/ by the phonological rule which nasalizes vowels before nasals does not change the interpretation of the word.

[14] It turns out that there is evidence for such markers in other languages.

(5.38) John is going./Is John going?

(5.39) Sally has taught linguistics./Has Sally taught linguistics?

The questions above are called 'Yes-No Questions' because (you guessed it) answers to such questions are either 'yes' or 'no.' (Contrast the answers to questions like 'What is he doing?') First, let's see if our PSR's will generate the above sentences. Right now, they allow for only a single V in the VP so if we think that 'is' and 'has' are V's, then the answer is definitely 'No' since 'going' and 'taught' are V's, too (adding up to two V's per sentence). In addition, our PSR's currently state the order for English as NP (subject) precedes VP (predicate). Both of the questions in the examples above, then, violate our PSR's since there is some element – the annoying 'is' or 'has' – which is not the NP subject coming in initial position in the sentence.

There are a number of logical possibilities for dealing with the 'is/have' problem but we will only entertain a couple of them here. First, we need to figure out what sort of syntactic category 'is' and 'have' belong to – do we, for example, need to add a new category or do they fit into an existing category? It seems as though V is the most likely candidate but even that is a poor fit. Some of the reasons which might make us vote against V are the following:

(5.40)

- John appears to be engaged in a 'going' activity rather than an 'is-ing' activity. Analogously, Sally was doing 'teaching', not 'has-ing.'
- Along the same vein, it is difficult to attach any kind of concrete meaning to 'is' and 'has', they seem to be more *functional* than anything else. Their most obvious function here seems to be to tell us something about tense (present and past, in these cases).
- Two V's together seem to be ungrammatical, in general – *John kicks hit(s) the ball.
- V's do not normally precede their subjects in English, under any condition – *Kicks John the ball.

None of these observations, taken alone, is compelling enough to argue that 'is' and 'has' are not V's. The combination, however, is probably sufficient (particularly if we note that we have greatly reduced the number and complexity of the possible observations on the topic) to force us to place 'is' and 'has' into a separate category from anything we have seen so far. We will use the traditional term for these elements, then, and call them 'Auxiliaries,' otherwise known, more informally, as 'helping verbs.' Now we have to decide how auxiliaries fit into to the existing syntactic relationships that we have posited – that is, how they fit into our PSR's.

PSR's model speakers' knowledge of possible syntactic relationships and structures. The current PSR's are based on a combination of speakers' intuitions and a variety of tests for constituency. For example, speakers feel that Adjectives 'go with' Nouns. But are there intuitions of this type about Auxiliaries? If we take just the two declarative sentences in the earlier example – 'John is going.' and 'Sally has taught Linguistics.' – we see that the Auxiliary (Aux) is adjacent to the NP(subject) and also to the VP(predicate). It could, of course, be a member of one of these two main constituents of the S, or it could be a member of *neither* constituent and instead be a separate (main) constituent. In fact, Aux's seem to have a relationship with *both* the NP and VP constituents. Notice that the Aux's must **agree** with the subject in number – if the subject is a third person singular, like 'John,' then any form of 'be' other than 'is' is ungrammatical.

(5.41) John is/*am/*are going.

At the same time, the Aux's appear to carry information about tense (i.e., when the verbal activity took place) – past, present, or future. In the end, then, we might think of Aux as being the constituent that 'ties together' the NP and the VP. It seems likely, in fact, that every NP and VP would *require* such a constituent. Without an Aux, we would have pairs of NP's and VP's like the one below.

(5.42) $[\text{John}]_{NP}$ $[\text{go}]_{VP}$

Here we have both an action (the V of the VP) and an entity who might perform the action (the N of the NP) so we have the critical components to predicate something of someone. What we *aren't* able to establish from just these two constituents is when the relationship between them took place – was it in the past, the present, or the future? It is the role of Aux in a

Sentence to establish such relationships.[15] Given this relationship to both the NP and VP, we can add Aux to our PSR's as a sister to *both* of them and rewrite our first PSR as follows.

(5.43) S → NP Aux VP

Now that we have a better idea of what we are working with, let's examine these pairs of sentences in further detail.

(5.44) John is going./Is John going?

(5.45) Sally has taught linguistics./Has Sally taught linguistics?

Our analysis of Aux has now enabled us to directly generate the first of each pair by means of our PSR's. Our PSR's will still not generate the interrogatives, however. Once again, we are faced with the issue of whether to revise our PSR's so that they will generate the interrogatives or to add a transformation to the syntactic component and derive the interrogatives that way. There are several things to think about whenever we are faced with such a question as the result of seeing new data. The first of these is whether the sentences are related to one another in some way. In the above examples, the only difference between the sentences is the force of the utterance – declarative or interrogative. The constituents themselves and the relationships between them remain constant. So, for example, that 'John' will be doing the 'going' is true in both the statement and the question. The second thing to consider is what will the effects of revising our PSR's be? If we were to revise our S rule such that it generated the questions above, we would have to state it as follows.

(5.46) S → (Aux) NP Aux VP

[15] These relationships of time are not the only ones that Aux's are responsible for, nor are 'be' and 'have' the only Aux's. Moreover, a quick glance around will show you that not all sentences in English contain obvious Aux's. As usual, the picture is more complex than we are able to explain here. Suffice it to say that our general statement about the necessity of an Aux holds true, even if it doesn't *seem* like the sentence has such an element in it. Because of this, we are going to make Aux an obligatory constituent of every S and simply note that it doesn't always have phonological content.

This rule says that a S *might* have an Aux in initial position and *must* have a NP Aux and VP (in that order). Recall that we want our PSR's to produce *all and only* the grammatical strings. If we modify our PSR's as above, will we still produce only grammatical strings? One way to determine the answer to this question is to slot appropriate lexical items into the PSR and exploit all of its possible options. Some of the strings that will be generated if we do this are given below.

(5.47)
- The boy is doing the dishes.
- *Is the boy is doing the dishes.
- Sally has written a letter.
- *Is Sally has written a letter.
- *Has Sally has written a letter.

As you can see, the revised S rule will produce a great many ungrammatical strings because it allows for *two* Aux's, one before the subject and one after it. This suggests that revising our S rule such that Aux's before NP's are directly generated is the wrong choice here.

Let's explore the other option – leaving our PSR's as is and producing the interrogatives by means of a transformation.[16] We would need a transformational rule which moves the Aux from its base-generated position (the PSR-generated position) to a position before the NP. This rule could be stated as follows:

(5.48) **Question Transformation**: Move the Aux to immediately precede the subject NP.

We would need to limit the application of such a rule in the same way we limited the application of the Imperative Transformation that we discussed earlier. We do not want *all* Aux's to move ahead of subjects, just those involved in Yes-No Questions. We also need to make sure that the input strings for declaratives and interrogatives are different because the

[16] Note that we are choosing to derive the interrogatives from the declaratives rather than the other way around. If we were to derive, instead the declaratives, we would have to revise our S rule to produce sentence-initial Aux's, as above. We have already seen that this type of revision will cause our S rule to over-generate (generate ungrammatical strings).

computational system produces different outcomes for each. As we did with imperatives, we will suggest it is necessary to include some special marker – a question marker – along with the lexical items in the string that is to be a question. This question marker will be the trigger for 'Question Transformation' to apply during the derivation, moving Aux to its surface or s-structure position.

(5.49)

d-structure [ $]_Q$ Sally has taught linguistics

$\Downarrow$

Q-trans

$\Downarrow$

s-structure $[\text{Has}]_Q$ Sally ___ taught linguistics

In addition to the fact that the 'revise PSR's ' alternative was ruled out due to over-generation, choosing the transformational rule analysis of pairs of sentences of these types has a number of advantages. The transformational analysis allows us to relate pairs of sentences that speakers' intuitions tell them are related by assigning all the elements of the strings exactly the same relationships at d-structure, the only difference being a marker for type of utterance. It also allows us to maintain a unitary analysis of Aux – there is one Aux category and Aux has the same relationship to both the NP and VP in every case.

English is not the only language which has this type of transformational rule for question formation. The examples from French and German below reveal a very similar process.[17]

(5.50) *Il a mangé du pain / A-t-il mangé du pain?*
he has eaten some bread / has-he eaten some bread
'He has eaten some bread' / 'Has he eaten some bread?'

(5.51) *Hans hat den Film gesehen / Hat Hans den Film*
Hans has the film seen / has Hans the film
*gesehen?*
seen
'Hans has seen the film' / 'Has Hans seen the film?'

On the other hand, some languages, like Telugu, form Yes-No Questions by adding a special question particle (the [ā] at the end of the sentence). It is possible to think of this question particle as corresponding roughly to the Q-marker we posited for the English question transformation, with the difference that the particle has some phonological content.

[17]We are glossing over some of the differences between French, German, and English here. The important point, for our purposes, is that Yes-No Questions in all three languages involve moving the Aux ahead of the subject.

(5.52)

| | | | | | | |
|---|---|---|---|---|---|---|
| *Ravi* | *guḍiki* | *veḷḷāḍu* | / | *Ravi* | *guḍiki* | *veḷḷāḍā* |
| Ravi | temple-to | went | / | Ravi | temp-to | went-Q |

'Ravi went to the temple' / 'Did Ravi go to the temple?'

While examples from other languages may seem to be dramatically different from English in many ways, such as the Telugu example above, it turns out that the differences are really rather minor. The same syntactic categories are present and the same sorts of syntactic relationships are established in both languages (and, in fact, in every human language).

## 5.10 Summary

We have seen that, in a general sense, the syntactic module of the grammar is similar to the phonological and morphological modules. In all three components, there is a more abstract level of representation (d-structure for syntax), a computational component (where, in syntax, transformations apply), and a less abstract representation (s-structure for syntax). A d-structure representation is formed by taking elements from the lexicon and combining them in a way that follows (some subset of) the PSR's, thus establishing certain types of constituents with particular relationships to one another. The interpretation of the string is based on both the lexical items and the relationships established between constituents. The d-structure representation is input to the computational component where transformations may be triggered by particular elements in the string. The computational system produces an output (the s-structure). As with the other modules, this process must be bi-directional. Our model of human syntactic knowledge must capture speakers' intuitions about constituency, ambiguity and relatedness, as well as phenomena we have observed such as recursiveness and creativity. As usual, we look for the most economical representations and operations possible which still account for the relevant data.

## 5.11 Exercises

1. Apinayé Syntax

   Examine the syntactic date from the Apinayé language below.

   (a) kukrẽ kokoi '(the) monkey eats'

(b) kukrẽ kra '(the) child eats'
(c) ape kra '(the) child works'
(d) kukrẽ kokoi ratʃ 'the big monkey eats'
(e) ape kra mɛtʃ 'the good child works'

a. What phrase structure rule for noun phrases (NP's) would you assume, from the data above, is part of a speaker of Apinayé's knowledge?

(a) NP $\rightarrow$ (D) (A) N (PP)
(b) NP $\rightarrow$ (A) N
(c) NP $\rightarrow$ N (A)
(d) NP $\rightarrow$ N

b. What phrase structure would you posit from the data above for Apinayé VP's?

(a) VP $\rightarrow$ V
(b) VP $\rightarrow$ V NP
(c) VP $\rightarrow$ V (NP)
(d) VP $\rightarrow$ V (NP) (PP)

c. What phrase structure rule seems to give rise to the structure of Apinayé S's?

(a) S $\rightarrow$ VP NP
(b) S $\rightarrow$ NP VP
(c) S $\rightarrow$ (NP) VP
(d) S $\rightarrow$ V NP

2. English Modals

Consider the following pairs of sentences. The modals 'can,' 'will,' 'should,' and 'could' can be classified as Aux's.

Mary can play the piano./Can Mary play the piano?

John will buy a new car./Will John buy a new car?

She should audition for the part./Should she audition for the part?

He could visit them more often./Could he visit them more often?

a. Our PSR's for English will generate:

(a) all of the sentences.
(b) only the statements.
(c) only the questions.
(d) none of the sentences.

b. Our PSR's for English in combination with the transformations we have seen for English will generate:

(a) all of the sentences.
(b) only the statements.
(c) only the question.
(d) only the sentences with 'will.'

# Chapter 6

# Semantics

## Contents


6.1 Introduction . . . . . . . . . . . . . . . . . . . . . . 134
6.2 Lexical Semantics . . . . . . . . . . . . . . . . . . . 135
6.3 Interpretation of Sentences . . . . . . . . . . . . 139
6.4 Summary . . . . . . . . . . . . . . . . . . . . . . . . 144
6.5 Exercises . . . . . . . . . . . . . . . . . . . . . . . . 144


## 6.1 Introduction

Until now, we have concentrated mainly upon the computational aspects of the linguistic system with little regard for the meanings of either words or utterances. However, the linguistic system *is* used for communication (even if that communication is only you talking to yourself) and therefore is capable of sending and receiving messages. Semantics is the study of meaning and the manner in which meaning is conveyed by the system. It is important to note that this does not include the study of *why* individuals choose to convey a particular meaning on a particular occasion. This ability lies outside the linguistic computational system and falls more properly under the fields of psychology and sociology, and perhaps several others.

An in-depth discussion of the intricate nature of formal semantics would require background knowledge of logic, philosophy, and possibly mathematics. Since this is beyond the scope of an introductory course, we will restrict ourselves to brief overviews of only two areas: lexical semantics and sentential semantics.

## 6.2 Lexical Semantics

As the name implies, Lexical Semantics is the study of word (or, more accurately, morpheme) meaning. We saw in Chapters 4 and 5 that lexical entries must contain certain information about morphemes. Part of this information is categorial – indicating whether the morpheme is a noun, a verb, an adjective, and so forth. Another part of this information might be termed 'functional' – for example, one morpheme might have the function of forming plurals, another might function to form third person singular presents of verbs. Yet another part of the information might be more closely associated with the everyday notion of 'meaning' as having to do with properties of real world objects such as 'dogs' and 'tables.' For our purposes, the greatest distinction appears to be between what linguists frequently term *content* morphemes, such as 'dog' and 'table' and *function* morphemes such as the plural and third person singular present markers. Because the semantics of content morphemes is generally easier to discuss (although by no means well understood), we will concentrate on members of that category.

We assume that the semantic system (like the phonological, morphological, and syntactic systems) contains units that are specific to it. It does not, for example, make reference to articulatory features as does the phonological system.[1] One of the goals of semantic inquiry, then, is to make a model of the type and set of features which the system operates on. One possiblity is that there is a (probably finite) set of semantic features available to human language. Every morpheme must be assigned some feature or combination of features from this set. How this might work is illustrated below.

[1]This is not meant to imply that the semantic system cannot 'see,' or make reference to, any aspects of the other systems. There is no doubt important interaction among all the systems involved in grammatical computations. This issue is independent of whether each system has units particular to it.

| | **girl** | **boy** | **baby** | **puppy** |
|---|---|---|---|---|
| | +human | +human | +human | -human |
| | +female | -female | | |
| (6.1) | -male | +male | | |
| | +young | +young | +young | +young |
| | -really little | -really little | +really little | +really little |
| | -4 legs | -4legs | -4legs | +4legs |

The above example is only meant to illustrate, in principle, how sets of semantic features could serve to distinguish morphemes. Note that the particular features above are almost certainly not good representations of whatever the true semantic features are. Actual semantic features are likely to be primitives – elements which are not able to be broken down further into constituent parts. Note that we have actually used interpretable morphemes and even combinations of morphemes ('human,' 'really little') as the semantic features to define other morphemes, such as 'boy' and 'baby.' [2]

Some of the features we have listed in 6.1 sound a bit implausible, particularly '+/-4 legs.' We could, of course, have used '+/-canine' and sounded a bit more sophisticated. We deliberately did not do so for two reasons. The first is that it is just as unlikely that an innate semantic feature system uses '+/-canine' as '+/-4 legs.' The 'canine' designation was developed by biologists and zoologists to classify a particular species (*Canidae*) according to certain anatomical (in this case, dental) features. It would be most surprising if this kind of classificatory system was something that humans were innately endowed with, particularly since it is, in this case, in English (and Latin). In fact, a very salient physical property like '4 legs' has considerably more appeal as a hypothesis. The second reason is that the unlikely-sounding '+/-4 legs' highlights the important distinction between a model and the thing which is being modeled. There is no *a priori* reason for any part of a model to either look like, or sound like, the thing it is modeling. Assuming that semantic features (or any other units and operations we posit for the human linguistic system) bear any resemblance to what is used by the real system itself is analogous to assuming that numbers fall down because gravitational law is represented by an equation.

In spite of the above difficulties with semantic features, there are still many interesting issues that we can consider. Note that in 6.1 we have not

[2]The fact that we typically use other words for features indicates something about our current understanding of lexical semantics.

really listed a sufficient number of features to distinguish a puppy from a kitten (or any other 4-legged 'infant'). The true feature system must be extremely detailed in order to capture such differences. On the other hand, there are certainly some morphemes whose semantic features are far less detailed. For example, the features for the morpheme 'male' (as in 'It's a male.') cannot be limited by a feature 'human' or by any feature which has to do with size or age. It seems, then, that there may be a superset-subset relationship between semantic feature bundles. It is also possible that these sets are organized hierarchically and are, therefore, directly related to one another. If such relationships exist, it is doubtful that they are similar to what we normally think of when we think of word relationships. Typically, when we think of how words are related to one another, we think in terms of things like synonyms (words that mean more or less the same thing) or antonyms (words that mean more or less the opposite of one another) or some other type of '-nym.'

Our traditional classification of word relationships, particularly synonyms, brings up an interesting question. Can any two morphemes have identical semantic feature bundles? That is, are any two morphemes truly synonymous? We believe that the answer to this question is 'no.' One of the reasons that we do not think true synonymy exists is because it would introduce a randomness into an operation that appears to be systematic. Whatever way the human brain translates some 'intention' or 'message' into a selection of morphemes and a syntactic representation, it seems as if it must do so based on some criteria. How would we choose between two morphemes with identical semantic feature bundles? Some choice must be able to be made on the basis of certain criteria. Two morphemes, both of which met the criteria, would not allow for a differentiating choice.

There is also the challenging question of how humans actually construct semantic feature bundles and associate them with some sequence of sounds. First of all, it seems as if this process has no end point. Adding to the lexicon continues throughout one's life. Second, one is able to make 'revisions' to lexical entries. It frequently happens that you thought a word had a particular meaning, found out that you were wrong, and subsequently altered the semantic features for that entry such that they were at least closer to what you inferred the word meant from the new information you received about it. Neither of these aspects appears to be present for adult speakers in the other parts of the grammar – in the phonology, morphology, or the syntax. As we will see in the next chapters, there is some point at which you are 'finished' constructing the rest of the grammar and after that point you cannot go back and revise it. New information received after that time leads to the construction of an additional grammar rather than modifications of an existing grammar.

How do children (and adults) form hypotheses about which semantic features should be associated with one another to make up a coherent representation? There is evidence from production and comprehension that children in the acquisition stage do not 'zero in' on an adult-like representations the first time around. Interestingly, children seem both to overgeneralize and undergeneralize. An example of the former is a child who refers to any number of (non-dog) animals as 'dog.' An example of the latter is a child who *only* refers to his or her own dog as 'dog.' There are a number of possible explanations for both of these phenomena, only some of which have anything to do with the accuracy of children's semantic representations. (We will leave these to you as a 'thought exercise' for now and return to it in question form at the end of the chapter.) In spite of these apparent false starts, the vast majority of children converge on something like the adult semantic representation of 'dog' as far as we know. In addition, however, they all perform what seem like magical leaps in the area of semantic representation. When they finally converge on 'dog,' they are able not only to identify virtually all and only the disparate members of the canine species as 'dogs' but also to identify stuffed animals that look like dogs as 'dog,' photographs, drawings and paintings of dogs as 'dog,' and things that bark in the night as 'dog.' Notice that the physical reality of stuffed toys, pictures, and noises are quite different both from one another and from a live dog. The evidence of 'doghood' must come from a sufficient number of semantic sources that one can remove a good many features and still be left with something that is 'dog.'

This very brief discussion of semantic features brings up the fundamental question of whether or not any two individuals have exactly the same featural representation for a particular morpheme in the same language, let alone for what we think of as an equivalent morpheme in some other language (cf. 'dog' and the French equivalent 'chien'). The assumption that there is a universal semantic feature set available does not entail the further assumption that everyone constructs identical semantic feature bundles. For content morphemes, experience must play a large role in what features are chosen to represent a particular morpheme. Since no two individuals have identical experiences, one might hypothesize that there exist a number of differences between semantic feature bundles across individuals. This is probably more likely in the case of abstract notions, such as 'like,' 'despair,' 'acquaintance,' and 'friend' than for concrete objects such as 'dog' or 'table' but may well be true of all content morphemes. However, if such differences exist, they seem not to impede communication to any significant degree.

It is possible that there is a critical difference between content morphemes and functional morphemes in the above respect. We have so far dealt only with a few content morphemes. We will turn now to functional morphemes and the role that these may play in the interpretation of sentences.

## 6.3 Interpretation of Sentences

If it is possible that no two individuals interpret the same content words in exactly the same way, is it possible that what appear to be the same sentences are also interpreted differently by different people? If the latter were true, it would seriously compound the difficulties in interpreting messages. In our model of syntactic structure, we assume that there are a number of syntactic categories and relationships that are language-independent – only trivial differences exist between surface word order across languages. We have made basically the same hypotheses for other modules of the grammar (phonology and morphology), that is, that there exists a set of universal features and operations available to human language. Our discussion of the semantic features for content words follows these assumptions in that we have posited a universal set of features. However, there is no guarantee that one individual's grammar will have the same semantic feature bundle as another individual's grammar for a particular lexical item. We would still like to believe, though, that the semantic component of the grammar processes input in a systematic fashion, as do the other components.

Note that semantic interpretations must involve more than just adding together the features of a set of content morphemes, otherwise strings like 'boy dog big tall' would be perfectly acceptable (i.e. interpretable) semantically. Additional information including which properties are being assigned to which entities and the force of the utterance (statement, question, command) must be supplied from other sources. Moreover, notions such as 'who's doing what to whom' and whether the utterance is a question or a statement appear to be somewhat independent of the interpretation of the content words in an utterance. Consider the following examples.

(6.2) a. The glor strued the zoobie at the blag

b. What zimz kaddle ording un ploget

c. Jore hof ap reeg trudle egdil

While 6.2 (a) and (b) are not stellar examples of clarity, they can be assigned some partial meaning. 6.2 (a) has the benefit of having what appears to be perfectly fine syntactic structure, which may increase the feeling that it is almost interpretable. One feels that if one only knew what a 'glor,' a 'zoobie,' and a 'blag' were and what sort of action 'struing' was, then the sentence would be perfectly clear. The syntactic structure of 6.2 (b) is not at all obvious. What *is* obvious about (b) is that it begins with 'What' and as soon as you read or hear 'what,' you are fairly certain that this is a question of some kind. In 6.2 (c), on the other hand, it is clear that we have just made up a bunch of stupid-sounding words and tossed them together on the page (no easy task...). There is no sense in (c) of even minimal interpretation possibilities.

We are going to discuss rather briefly the nature of words like 'what' and the kind of role such elements may play in the interpretation of utterances. Our basic assumptions will be the following. The first is that the semantic component will produce, for a given set of features, the same output (interpretation) every time and that this is true not only within individual grammars but for all grammars. The second is that there is some set of semantic units which are functional elements that *all* humans possess innately, some or all of which are active in all adult grammars. For instance, all languages have questions, as well as statements, and we assume that there is some semantic feature or features (possibly very abstract) which mark these two types of utterances for the different interpretations. Thirdly, we believe is is possible, in principle, to separate the process of interpreting the formal properties of utterances based on functional elements from the semantic features of the content words involved in those utterances. (This was, at least in part, the exercise involved in 6.2).

We are hopeful that examining WH words will help to illustrate the points we have made above. In English, WH words are all of the interrogative pronouns, most of which begin with the sequence <wh...> and one of which begins with just <h>, i.e., <how>.[3] All languages have some such set (not

[3] There are other words which look and sound just like these but do not have the same function or distribution, such as the set of relative pronouns. We are only interested in the interrogative pronouns here.

necessarily words) and we conveniently refer to all such sets as WH words, independent of the language in question (and all phonological evidence to the contrary). There is a noticeable pattern in the placement of WH words in English – if there is only one WH word, it always occurs first in the sentence.[4]

(6.3)
a. What did the child see?
b. Who did the child see?
c. Where did the child go?
d. Why did the child go?
e. How did the child fall?
f. Which child did Susie see?
g. *Did the child see what?
h. *Did the child see who?
i. *Did the child go where?
j. *Did the child go why?
k. *Did the child go how?
l. *Child did Susie see which?

You may be able to think of apparent counterexamples to the first set (3a-f) above. (The ungrammatical examples in 6.3 are there for comparison purposes only and we won't refer to them again.) For instance, it is perfectly grammatical to say 'The child saw **what**?' This type of question is known as an 'echo question.' Echo questions have a different interpretation than the acceptable types of WH questions in 6.3 and are used in different contexts. Note that 'What did the child see?' is a request for new information in the sense that the listener has not yet heard what it is that the child saw. On the other hand, the echo question 'The child saw **what**?' is a request to repeat information that has already been presented but which, for some reason (usually either mishearing or disbelief), the listener did not process.

(6.4)
- The child saw a Martian.
- The child saw **what**?

[4]The actual data on WH questions is far more complex than these examples suggest but this subset will be sufficient to illustrate our point.

Echo questions also have a different intonational pattern than their simple WH counterparts. The placement of the WH word in echo questions is relevant to our discussion, as well, and we will come back to it later on. For the moment, we will concentrate on the simple WH questions in 6.3 .

It is mildly curious that all of the Wh words in 6.3 seem to have to be at the front of the sentence. (We know that it is unlikely to be due to some phonological property of the sound sequences of those words because they *can* occur elsewhere, as echo questions show.) It is even more curious when we look at WH questions in other languages. Many languages have WH questions which are like those in the Telugu examples, below.

(6.5) *Pilla ēmiṯi tʃūsindi*
child what saw
'What did the child see?

(6.6) *Pilla ēvaru tʃūsindi*
Pilla who saw
'Who did the child see?'

In Telugu, the WH words do not all occur at the front of the simple question. In fact, their distribution is instead just like the WH words in echo questions in English, allowing for the difference in Telugu word order. Critically, however, the interpretation of the Telugu simple questions above is *not* as echo questions but is identical to the English 'What did the child see?' and 'Who did the child see?', respectively.[5]

The critical question to ask at this point is whether we are counting apples and oranges (i.e. things that cannot be counted together because they are completely unrelated to one another) or whether we have, at some very abstract level, essentially an utterance with the same function and force in both English and Telugu. Certain pieces of information may be relevant in answering this question. First, the similarity between the placement of the WH word in English echo questions and Telugu WH questions. Second, children at the acquisition stage in an English environment do not produce strings such as 6.3 (g-l), nor any strings which have the WH word other than at the front or in the echo question position. Third, speakers of Telugu (and languages like Telugu with respect to WH questions) find the 'switch' to

[5]Telugu does have echo questions but they are indicated by a postposition which is not present in these examples.

English WH questions completely trivial. Fourth, Telugu-English bilinguals accept the Telugu questions and their English equivalents in **??** as being 'the same.'

This is hardly a scientific approach to the question. The information that we have asked you to consider in the preceding paragraph is largely anecdotal and not substantiated with empirical evidence. However, a more careful consideration is well beyond the scope of an introductory text. As it is, we can only say that all of these things are suggestive of a hypothesis which treats the simple English WH questions and their Telugu equivalents as being the same at some level of interpretation. If we adopt such a hypothesis, we must still account for the fact that the surface forms are different–somehow we have to make two strings which appear different be the same.

If you recall, in our discussion of allophones and phonemes in Chapter 3, we stated that allophones are just surface representations of an underlying mental representation, the phoneme. The case of WH words might benefit from a similar (though not identical) analysis. One possibility is that the features of a WH word such as 'what' are relevant to two aspects of interpretation. One of these is that 'what' is a 'stand-in' for some NP. In this sense, 'what' is 'NP-like' and needs to be in a position that NP's can be in. This requirement would be satisfied by the position of 'what' in the Telugu sentence or in the English echo question case. In both of these cases, 'what' can be replaced by an NP, such as 'the dog'. At the same time, however, 'what' appears to have a feature that causes an utterance to be a question instead of a declarative. Note that the question feature seems to affect the entire utterance, not just one word. This can be compared to the English echo question where the question feature of 'what' *does* seem to affect just one word. It is possible that, in order to have an effect over the entire utterance, 'what' may need to be at the beginning (or perhaps at the very end) rather than in the middle somewhere. From a syntactic point of view, this would place it 'over' the rest of the sentence in hierarchical structure terms. It seems, then, as if 'what' needs to be in two places at once or perhaps have two copies, one for each position. If the distinction between English and Telugu is truly only a superficial one, then it might simply be that both languages require a copy of 'what' at the beginning and both require a copy of 'what' in an NP position and English pronounces the 'what' at the beginning whereas Telugu pronounces the 'what' in the NP position. At an abstract, non-phonetic level then, English and Telugu will be the same (and, by hypothesis, any case of a 'what' question in any other language as well), it is only a matter of which 'phonetic' choice is made (based on some assumption that phonetic redundancy is avoided) that causes them to appear different.

Note that the task of acquiring language will be much less difficult under a scenario where this type of variation across languages is just a superficial property. This is particularly true of the components of the grammar which are inherently more abstract, like syntax and semantics. Correspondingly, one might assume that these more abstract components are innately more specified than other components for the simple reason that it would otherwise be considerably more difficult to deduce their operations from surface structure.

## 6.4 Summary

The level at which meaning exists in language is undoubtedly the most abstract and thus the least amenable to an empirical approach. Our very brief discussion of it here is intended primarily to offer support for the view that universal properties of language exist in semantics as they do in the other parts of the grammar that we have discussed. In addition, it seems that semantic units (in particular 'functional' units) and computations are a critical part of constructing and interpreting sentences, respectively.

## 6.5 Exercises

1. If a child in the early stages of acquistion refers only to his/her dog as 'dog', what might this indicate?

   (a) The child's grammar takes 'Spot' and converts it to 'dog' by regular, phonological rules.

   (b) The child has taken 'dog' to refer to the very specific set of attributes of his/her particular dog (i.e., hair of a particular length and color, bulgy brown eyes, wart on the chin, and so on).

   (c) The child will never converge on some adult-like meaning of 'dog.'

   (d) The child initially assumes that the sound [d] refers to hairy objects, and [g] refers to 4 legged objects.

2. Which of the following is true?

(a) Some morphemes have meaning and some do not.
(b) Some morphemes have 'content' semantics and some 'functional' semantics.
(c) All morphemes have both 'content' semantics and 'functional' semantics.
(d) The semantics of simple morphemes are derived during the course of a derivation.

3. If a child in the early stages of acquisition refers to a bunch of 'non-dog' animals as 'dog,' this indicates that:

   (a) the child's knowledge of language includes only phonological features at this point, not semantic features.
   (b) the child has heard everyone else refer to those other animals as 'dogs.'
   (c) the child cannot distinguish between 'non-dog' animals and dogs.
   (d) the child has overgeneralized – later he/she will refine their semantic feature set for 'dog' to be more specific.

# Chapter 7

# Sociolinguistics

## Contents

7.1 Introduction . . . . . . . . . . . . . . . . . . . . . . 146
7.2 Traditional Dialectology . . . . . . . . . . . . . . . 148
7.3 Problems with Traditional Dialectology . . . . . 151
7.4 The Sociolinguistic Revolution . . . . . . . . . . 152
7.5 Standard and Non-Standard . . . . . . . . . . . . 156
7.6 Gender and Linguistic Variation . . . . . . . . . 158
7.7 Summary . . . . . . . . . . . . . . . . . . . . . . . . 160
7.8 Exercises . . . . . . . . . . . . . . . . . . . . . . . . 162

## 7.1 Introduction

Up to this point in the course, we have been concerned almost exclusively with the nature and structure of the grammar – the linguistic 'knowledge' that resides in the mind/brain of an individual speaker. We have discussed the various types of knowledge that appear to be involved (phonological, morphological, syntactic, etc.) and posited 'modules' of the grammar to account for how these different aspects of linguistic knowledge appear to function. For the rest of the course we will focus on several interesting questions which relate to these individual knowledge states, including:

- Why, if such knowledge is a property of individuals, do people from the same area, class, gender, ethnicity, etc. seem more similar to one another than to others with respect to this knowledge? How does it come about that this is the case?

- Why, given that such knowledge is a property of individuals, does it appear (from our historical records) that people who spoke English-type grammars five hundred or a thousand years ago had such different knowledge states than those who have English-type grammars do today? Why does no one speak 'Old English' or 'Middle English' today? What do these terms mean?

- What does it mean to say that two 'languages' are related? Does it mean anything, now that we recognize that 'language' in the traditional sense is not a very meaningful scientific concept?

- How did the knowledge state that an individual has in his or her brain/mind get there? What is the learning mechanism? How much of what we know, when we know a language, is learned and how much is part of our genetic inheritance as humans?

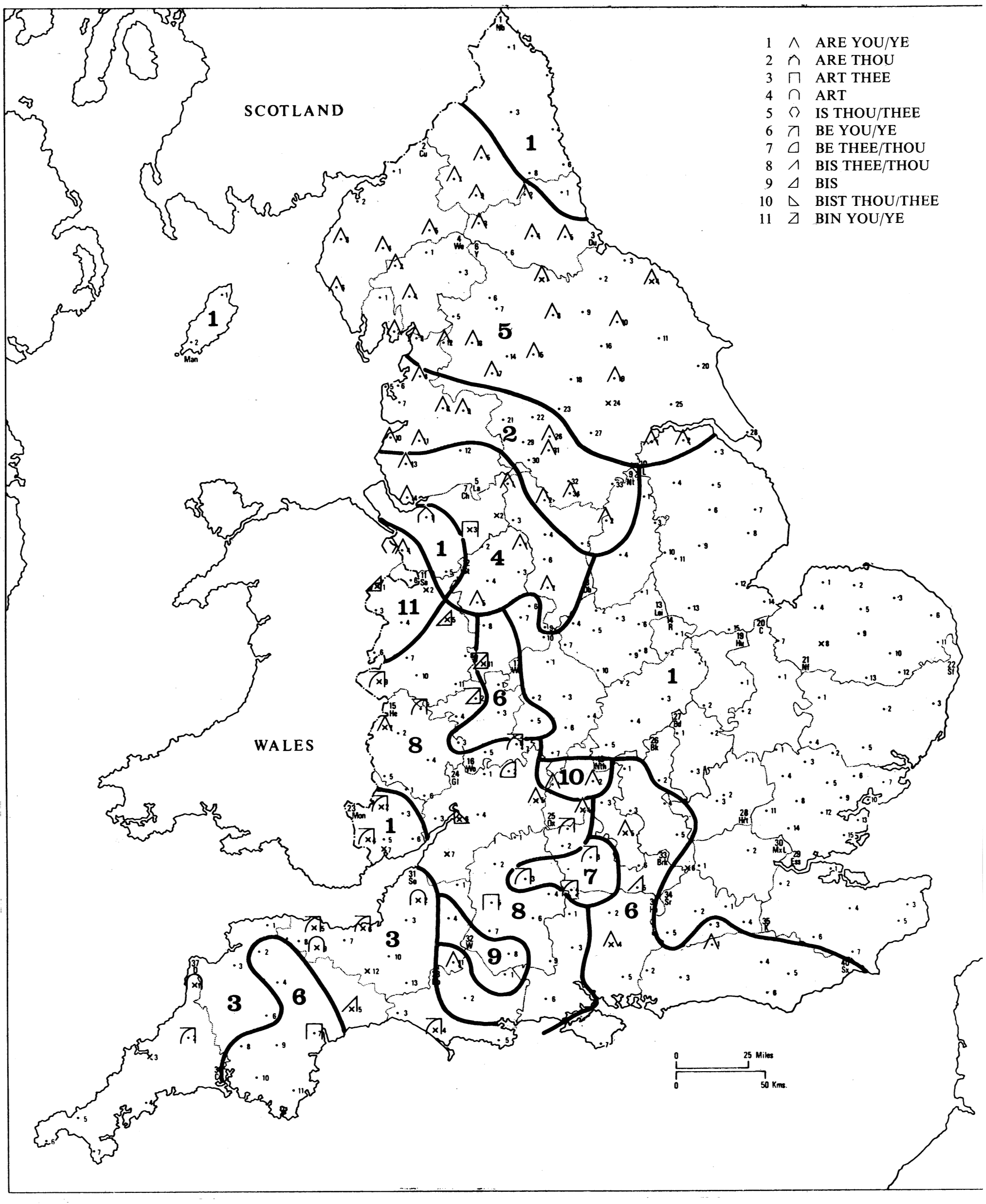
The first set of questions forms the central focus of the field known as 'sociolinguistics,' which will be our primary concern in this chapter. The second set of questions forms the core of the study of historical, or diachronic, linguistics which will be treated in the first part of Chapter 8. The third set of questions, those regarding linguistic classification, are the subject matter of comparative linguistics and will be dealt with in the second part of Chapter 8. Finally, the fourth set of questions represents the central focus of the field of language acquisition, the subject matter of Chapter 9. The study of language acquisition allows us to see what contribution the field of linguistics can make to the important general philosophical debate on the so-called 'nature' vs. 'nurture' issue. We hope it will become clear to you in the course of your reading that the sets of questions given above are in fact intimitely related to one another – in spite of their apparent diversity.

## 7.2 Traditional Dialectology

As any of you can doubtless attest from your own experience, the precise form of individual grammars is not uniform – different people speak differently. It is of some interest to note that the differences that we observe in the way people speak are not randomly distributed across the population. Instead, speakers from similar geographic areas, for example, tend to speak in a similar manner. At the same time, however, speakers from roughly the same geographic area (e.g., Boston) may still sound quite different from one another. These differences often depend upon such factors as social 'class,' age, ethnicity, gender, religion, occupation, etc. The study of the relationship between 'social' factors, broadly construed, and linguistic variables (for example, whether or not one pronounces the [r] of words like 'farm') is known as *sociolinguistics*. The study of the relationship between geography and linguistic variables is generally known as *dialectology*. In this chapter we will consider the nature and major results of these two disciplines.

The fact that people spoke somewhat differently in different places was apparent to linguists (and probably to everyone else who thought about it for a moment) from early on. In the 19th century the primary focus of linguistic research was *historical linguistics* – i.e., the study of the relationships between languages and how these languages came to look the way they did. 'Language' was of course not interpreted in the modern theoretical sense (in which 'language' equals 'grammar') and thus a certain amount of effort was invested in determining the answer to questions such as 'Is Dutch a different language from German, or a dialect of the same language?' We recognize the futile nature of such questions today, but in an attempt to deal with such matters in the 19th century an investigation of geographical variation in the relatively local sense – i.e., variation within a single 'language' (socio-politically conceived) – was begun. It was quickly discovered that the study of dialects in use in particular localities was of considerable interest to linguists' primary questions of the day (remember, these were historical questions, essentially). As an example, take a look at the following page (taken from the *Linguistic Atlas of England*).

ARE YOU (married) M2(




AREN YOU 11Sa 2
BE 31So 12, 36Co 3
BIST 24Gl +5/7
IST THOU 6Y +1/24

This map divides England into various areas based on the way people in these areas say the verb and pronoun in the question 'Are you married?' As you can see, responses to the inquiries of dialectologists on this matter were quite diverse. The lines which divide the area studied into 'dialect regions' within which some particular form (or, in this case, set of forms) predominates are known as *isoglosses*.

How are we to interpret the 'isoglosses' found on such dialect maps? Take another example. As one drives south from Michigan to southern Indiana, one crosses a major midwestern American isogloss, the so-called *pin/pen-line*. South of this line, the vowels of 'pin' and 'pen' are identical (because of a historical phonological event whereby the ɛ of 'pen' raised to ɪ before nasals), while north of the line the vowels of these (and phonologically similar) words are distinct. Let us get some basic facts straight first. It is not the case that, as you cross this line, the vowels of 'pin' and 'pen' suddenly merge in your own speech (if they hadn't before) – i.e., the isogloss doesn't control the behavior of humans, it doesn't *cause* the phenomenon it attempts to describe. So the isogloss clearly does not mean 'everyone south of this line merges 'pin' and 'pen' and everyone north of it keeps them distinct.' Instead, isoglosses usually[1] represent the mid-point in the distribution of a variable. Excluding people temporarily present on a particular side of the isogloss – i.e., considering only long-term residents, born where they now reside – in general we will find a point north of the pin/pen-line where everyone keeps the vowels of these two words distinct. At a point some distance south of the isogloss, we will reach a point at which all lifetime residents of the area show in their speech merger of the vowels of 'pin' and 'pen.' The isogloss represents the midpoint between these two extremes, and of course we generally find, as we approach the line, that the distribution of people showing or not showing the merger in question approaches 50-50.

Isoglosses, as we saw from the first example, are not restricted to phonological variation. They can separate speakers by morphological, syntactic, or lexical features as well. For example, as one approaches the pin/pen-line, one also runs into the famous bag/sack-line. North of the bag/sack-line a little paper container, open on the top, is called a 'bag.' South of it, this same object is called a 'sack.' That is, the same object has two different names in the two regions. Lexical variation is widespread – many of the

[1] We exclude fairly uninteresting cases in which the isogloss represents an actual major physical barrier, like the Atlantic ocean.

differences between American and British English involve their respective lexicons ('trunk' vs. 'boot;' 'cookie' vs. 'biscuit'; 'truck' vs. 'lorry'; and so on) although differences in many other areas exist as well.

## 7.3 Problems with Traditional Dialectology

Since traditional dialectologists were in the business of providing evidence for scholars who were working specifically on historical linguistic themes, it was most valuable for them to find, as they went from village to village, the oldest informants they could. After all, someone who is 90 years old learned their first language 88 years ago or so, while someone who is twenty only learned the language 18 years ago. If one is interested in history, the 90 year old informant thus has things to offer that a 20 year old cannot. Moreover, it was important that the informants selected by the dialectologist be long-term residents, preferably from birth, of the village they were being taken to represent. Neither of us (authors) would be a good choice of informant, for example, since neither of us currently resides in our place of origin and, as yet, neither of us has adopted the manner of speaking of those currently around us. In addition, if you could find someone who had basically never talked to anybody outside of the village, you were likely to get the 'purest' data for how that village spoke long ago, since you have minimized the possibilities of linguistic contamination from outside the village.

These two factors – age and long-term residency – had, as an effect, that dialectologists selected old, lifetime village residents who had had little contact with anybody outside the village. This meant, given life in the modern world, that the person selected to represent a given community on a dialect was not actually very representative of the village itself. That is, if you went to the village, you would not necessarily find that the average villager (who wasn't so old and maybe travelled around a bit...) talked the way the dialect atlas indicated the residents of the village would.

In addition, in communities which showed internal diversity along some social dimensions (note that this includes virtually every community), such as class or ethnicity, it was difficult to pick an individual who might be representative of the group as a whole. Everyone already knew, as we know now, that people from different levels of social structure talk differently, the working classes talking quite distinctly from the upper classes. In addition, it was known that if there was, for example, a sizable black community in a given town, that it was likely that the members of this community would not talk like the non-black residents. In keeping with the somewhat classist and racist assumptions of social science at the time, dialectologists generally sought middle class whites as 'representative' of the community under study. This skewing in the informant selection process had, needless to say, a corresponding effect on the accuracy of the dialect to geographic area mapping.

In cases of extreme internal racial, ethnic, and class diversity – as with most major cities in North America, for example – it was felt that no individual could represent the community. This problem was dealt with in a relatively straightforward, if decidedly non-scientific manner: cities were skipped. So the fact that each neighborhood in London, for instance, has its own dialect is nowhere reflected in the *Linguistic Atlas of England.* The same decision was made regarding major North American cities.

## 7.4 The Sociolinguistic Revolution

In keeping with the increasing urbanization of the population in the late 1950's and early 1960's, and with the demand for 'social relevance' of science that accompanied major social upheavals such as the civil rights and women's liberation movements in North America, the shortcomings of traditional dialectology came to be felt as fatal flaws to the enterprise. Questions that traditional dialectology, because of the nature of its data collection techniques (arising in part from its goal as servant to the historical linguistic enterprise), simply couldn't ask came to be of major concern. What do people in this place actually talk like (not just the old, kind of quiet and immobile ones)? What happens – linguistically – when someone moves across an isogloss? What relationship does linguistic variation have to such social variables as ethnicity, age, gender, class, etc.?

Clearly a different approach to the study of linguistic variation was required if we were to make some progress on understanding these issues. Primarily through the efforts of a single dedicated scholar – William Labov, now at Penn – the field of *sociolinguistics*, which attempts to understand the relationship between social variables and linguistic ones, began to flourish.

Labov is perhaps most famous for his well-developed sense of experimental design (he was originally trained as a chemical engineer, which may have helped). For example, he was interested in getting federal funding for a fullscale investigation of the sociolinguistics of New York's Lower East Side. As you may know, to get federal funding for research you generally have to do a 'preliminary study' to show that there is a good likelihood that your research, if funded, is going to be fruitful. Unfortunately, Labov had no money to undertake such a study. In spite of this, he devised a research project which, with no money and only his own effort, would yield interesting and valuable results. This now famous study is known as the New York Department Store Study.

The project was simple in its conception. Labov took three socially-stratified department stores on the Lower East Side: Saks Fifth Avenue (an upper to upper-middle class store), Macy's (a middle to upper-working class store) and Klein's (like a K-Mart – an essentially working class store).[2] He followed the same procedure for each of the three stores. He went up to the fourth floor to see what was there. Whatever goods were there on the fourth floor formed the basis of his survey questions for people in that particular store. He then proceeded to other floors, going up to people and asking, for example, 'where are the women's dresses?' The person asked would respond (if they knew) 'the fourth floor.' This gave Labov two chances to note down whether or not postvocalic /r/ was being pronounced by people in these various stores – postvocalic /r/ was the linguistic variable he was interested in.

The stroke of brilliance in Labov's study was the following. After getting a response to his question which indicated that the person he had asked knew where the women's dresses (or whatever) were, he said 'excuse me?' This conveyed to the speaker that the message had not gotten through, and so they repeated it, being more careful, slower, and more attentive to the act of speaking. For each person he approached in this way, then, he ended up with four instances of postvocalic /r/: two 'casual' and two 'careful' realizations.

One of the tables showing some of the results of this study is printed in Figure 7.1. This table reveals how the realization of /r/ (it's presence in both 'fourth' and 'floor') correlates with Labov's estimate of the age of the

[2] Labov walked through the stores after he had collected his linguistic data establishing empirical tests for the class differentiation. Some of these were physical aspects such as how much merchandise was out where the customer could see and touch it (in Klein's all of it was – the place was pretty junky-looking, in Macy's most of it was, in Saks one generally had to tell a salesperson what one was interested in and they brought out samples). There was a lot of open space to mill about in Saks's, but in Klein's one wandered through narrow isles through stacks and stacks of goods for sale.

| estimated age | Saks (Upper Middle) | Macy's (Lower Middle) | Klein's (Working) |
|---|---|---|---|
| 15-30 | 67 | 21 | 10 |
| 35-50 | 26 | 26 | 0 |
| 55-70 | 13 | 39 | 4 |

Figure 7.1: Percentage of speakers with two /r/'s in casual 'fourth floor.'

speaker.

The results are quite interesting, both because they provide us with a snapshot of how people from various classes actually spoke in day-to-day interaction on New York's Lower East Side in 1966, and because they give us some idea of the correlation between the robust presence of postvocalic /r/, social class, and age in that community. As you can see from the chart, older upper middle class speakers realize /r/ in percentages more like that of the youngest group of working class speakers – while the youngest group of upper class speakers are as far away from the youngest group of working class speakers as anyone. How can we explain this?

As is apparent if you watch American movies from the 1930's and 1940's (most of which were shot in New York) and contrast the speech heard in those movies with that we take as characteristic of New York today, the status of so-called r-lessness has changed considerably during this period. In general, the well-to-do speakers of the movies of the 30's and 40's are r-less, while the well-to-do of today's New York are r-ful. In addition, we know from historical records regarding educational policy in New York that r-lessness was explicitly taught as the norm in New York schools until just after the Second World War (ended 1945). Older upper middle class New Yorkers in Labov's study spoke in keeping with the sociolinguistic situation which was in place when they learned to speak: that is, they were generally r-less. Younger upper middle class speakers acted in accordance with the current sociolinguistic situation: they were generally r-ful.

Another interesting result of Labov's study, not clearly indicated by Figure 7.1, is that there were differences in some individual's realization of postvocalic /r/ in casual vs. formal speech samples. Specifically, *if* there was a difference (for a particular speaker), it was always the case that the casual sample was r-less and the formal sample was r-ful. A sensible question to ask at this point would be 'How can one grammar produce both r-less and r-ful output?' We hope that this query was inspired by the fact that you know that phonological rules cannot be turned on and off at will. The answer to the question depends upon one's view of what's going on with r-lessness. Under one scenario, r-lessness *is* produced by phonological rule. In this case, there is only one possible answer to the question, in our opinion, and that is that the individual is bidialectal. This is, of course, identical to being bilingual. Since there is no way to define 'language' in a linguistic framework and also no way to define 'dialect,' the terms bilingual and bidialectal reduce to the same thing.

The second possible scenario involves not the grammar but the lexicon, along with some type of 'conscious' cognitive processing. This second scenario develops as follows. You are exposed to another dialect. You notice, after a little while, not only that they talk funny but that there seem to be some consistent differences in phonology, morphology, and syntax between the way you talk and the way they talk. This is usually realized by thinking something like 'When I say X, they say Y.' Critically, these realizations proceed on a case-by-case basis initially (and often permanently). So, for example, one of the authors of this book is not from Michigan. This one has noticed that when she would say 'I have run up and down stairs all day.' many people from Michigan would instead say 'I have ran up and down stairs all day.' She has filed this information away as a single case of correlation between something she says and something 'they' say. It is possible (though unlikely) that this and other forms like it are produced by rule in the grammar of many Michiganders. If so, there is a chance that this author will, given sufficient time, exposure, and incentive, become bidialectal and will internalize this rule along with many others. On the other hand, it is possible (and probable) that this difference is the result of the author's having a list of irregular past participle forms in her lexicon which does not match the list of irregular past participle forms in the lexica of many Michiganders. If this is true, it is highly unlikely that this author will ever be able to produce the Michigan forms with any consistency. To do so would require that the author hear and consciously remember the Michigan correlate for every past participle on her list and, for each utterance, consciously *translate* her output into the correct Michigan form. A process like this correlation one described has nothing to do with the systematic workings of the grammar. The output of correlating and translating will be slower than normal and also riddled with odd, performance-related mistakes.

One type of mistake that this ‘correlation technique’ frequently results in is something that has been termed *hypercorrection*. Hypercorrection is the result of some general awareness on the part of a speaker that there exists a correlation between certain things they say and certain things speakers of another dialect say but the speaker is not too sure of the particulars. For example, a characteristic of Brooklyn dialects was the [bojd] pronunciation for [br̩d], that is, in many lexical items where the standard had a syllabic [r̩], Brooklyn dialect had the diphthong [oj]. The attempts on the parts of those Brooklyn speakers who were *not* bidialectal to produce the standard resulted in forms such as [kr̩l] for <coil>, [dʒr̩n] for <join>, and [tr̩lɪt] for <toilet>. It is easy to see how such a situation could come about. A Brooklyn dialect speaker was aware of some [oj]-[r̩] difference between their dialect and the standard. Unfortunately, the lexicon of the Brooklyn speaker included both words had [oj] in them for *all* speakers (such as the three above) and words that had [oj] in them for only Brooklyn speakers (e.g., [bojd]) due to some historical change that affected only that community. The problem facing the Brooklyn speaker was ‘Which words do I translate into [r̩] forms?’ In these sticky cases, speakers often opt for the ‘overdoing it is better’ strategy and end up ‘hypercorrecting’ – ‘fixing’ forms that do not need ‘fixing.’

It is difficult to determine exactly which of the scenarios described applied to the individuals in Labov’s study who showed both r-fulness and r-lessness. Most probably, some of the respondents were bidialectal and others were simply ‘translating’ (more or less successfully). We have not yet discussed the reason why, in the case of someone producing both forms, it was systematically a change to r-fulness from r-lessness in the careful speech forms. We will turn to that question and related issues in our discussion of standard and non-standard forms below.

## 7.5 Standard and Non-Standard

When we described the Broolyn dialect feature earlier, we used the term ‘standard’ for the non-Brooklyn comparison dialect. So just what *is* the ‘standard’ and where did it come from? Well, it is not something defined with respect to Brooklyn, New York, such that everyone who lives outside that borough (but no one who lives in the borough itself) speaks the standard. The very notion ‘standard’ is predicated on the non-linguistic assumptions that there is some grammar plus lexicon that the majority of people share (this relates to the usual idea that whatever most people do is the 'common' or 'standard' thing to do) and that, by virtue of some completely independent social metric (independent of the sheer numbers formula cited above), one 'dialect' is to be held as the 'purest' incarnation of the language. It is easy to see that neither of these assumptions is going to lead to a single, definable entity that can be, for example, 'standard English' and, indeed, no such entity actually exists. First of all, no two individuals share exactly the same knowledge state with respect to language. Prescriptive grammars and the educational system attempt to remedy this situation, in part, but they focus on only a tiny subset of natural language features for 'correction' (and are, in any event, consistently unsuccessful). As far as the 'purity' goal is concerned, it is, of course, linguistically impossible to evaluate different forms for 'purity.' Moreover, even a non-linguistic approach is going to run into logical problems with the idea that some forms are better or more pure than others since what is considered better or more pure seems to be quite changeable over time and situation.

It turns out that describing what is 'standard' involves talking about where a standard comes from. At any particular point in history, the standard tends to be the dialect of the social group which at that time possesses prestige, socio-economic status, and power relative to other social groups. Note that it may not be enough simply to have enormous prestige, for example, because money and sheer numbers in the right places are important, too. The standard for British English (called Received Pronunciation (RP)) is not the dialect of Her Majesty Queen Elizabeth II, but rather the dialect of a group with slightly less prestige (in terms of social class). (RP is typically exemplified by BBC broadcasters.) In the US, the current standard was drawn from the dialects of white, upper middle class people who were from other than Atlantic coastal or southern areas (broadly defined). Interestingly, the US model seems to be the US counterpart of a BBC broadcaster, Walter Cronkhite. This is still, of course, very vague. On occasion, it is easier to rule out forms than to describe what might be included in anything that is called 'standard.' Even here, the ground is murky. 'Non-standard' forms come in two types – those that are *stigmatized* and those are are 'non-stigmatized.' Stigmatized dialect features suffer from guilt by association. The existence of stratification within society, while dependent upon many different variables, means that some social groups are on the bottom relative to others. These groups tend to be stigmatized and, as a result, the dialect features particular to the group tend to be stigmatized, as well. After some time, the direction of causality is lost and anyone who has the stigmatized dialect features is, him/herself, stigmatized regardless of whether they happen to be a member of the stigmatized social group. This is completely illogical but probably no more so than having social stratification in the first place. On the other hand, 'non-stigmatized' forms which are 'non-standard' may be noticed but used with impunity (socially speaking). The fact that, in some dialects (including the current standard), the word for a cooking vessel is [pat] while in other dialects it is [pɔt] is unremarkable. In cases such as these, the forms which are, strictly speaking, non-standard will be accepted as standard (just a 'different' standard).

The existence of a (somewhat normative) standard is another facet of the interaction of social variables and linguistic features. The majority of sociolinguistic studies focus upon standard and non-standard (frequently stigmatized) linguistic features with respect to various social variables. The 'direction' of correlation is most often similar to that of the Department Store study where people tend to 'switch up' (toward the standard) if they can. In the following section, however, we will see that this is not always the case.

## 7.6 Gender and Linguistic Variation

Another study which showed some correlation between linguistic forms and social class was conducted by Trudgill (1996). In this study, five different groups in Norwich, England were examined for the production of [ɪn] (a non-standard, somewhat stigmatized form) and of [ɪŋ] (possible final sounds in such words as 'running,' 'jumping,' and so on). The five groups were: Lower Working Class (LWC), Middle Working Class (MWC), Upper Working Class (UWC), Lower Middle Class (LMC), and Middle Middle Class (MMC). Data from both casual speech and reading (i.e. very careful speech) were collected.

| Group | Casual | Reading |
|---|---|---|
| LWC | 0 | 34 |
| MWC | 5 | 56 |
| UWC | 13 | 85 |
| LMC | 58 | 90 |
| MMC | 72 | 100 |

Figure 7.2: Percentage [ɪŋ] forms produced.

The results indicate that speakers of higher socio-economic classes (LMC and MMC) produce a much greater percentage of [ɪŋ] forms in casual speech than do speakers of lower socio-economic classes. Speakers in *all* groups shifted 'upward' (toward the standard form) in careful speech contexts with the most dramatic shift being found in the UWC group. Interestingly, it appears to be characteristic of the UWC (and the LMC in surveys of US informants) that they show this type of greater shift toward the standard. One hypothesis put forth to account for this is that these social groups suffer from higher levels of 'social status insecurity' and therefore have a greater incentive to emulate linguistic variables associated with social groups above them in the social ladder.

The data in the Norwich study might have illustrated an additional relationship between linguistic variables and social variables if more finely tuned social variables were noted – particularly, the gender of the informants.[3] In culture after culture, we observe that women consistently speak like men of the next higher socio-economic class. Moreover, when asked about their own behavior (e.g., if asked 'how often to you say [ɪn] instead of [ɪŋ] in words like *running, jumping, playing*?) women consistently underreport stigmatized forms (like [ɪn]), claiming to use such forms only rarely, while men tend to *overreport* the use of stigmatized forms, claiming to say [ɪn] far more often than observations of their behavior in informal situations actually indicates is the case. These gender behavior differences (in terms of forms actually used, not forms reported by respondents to be used) manifest themselves well before individuals reach adulthood. In a study of the casual speech of twelve boys and twelve girls aged three to ten in a New England village,

[3]Note that, unlike some other social variables – age, income or occupation, etc. – gender is essentially immutable.

Fischer (1958) noted the following results.

| | [ɪŋ] | [ɪn] |
|---|---|---|
| Girls | 10 | 2 |
| Boys | 5 | 7 |

Figure 7.3: Number [ɪŋ] and [ɪn] Forms Produced By Boys and Girls Aged 3 to 10.

The explanation for this phenomenon is not entirely clear. The role of women as the primary bearers of responsibility for 'social propriety' has been suggested, as has the general riskiness of their social status relative to men in the same social class. Another suggestion is that there is a 'masculinity factor' associated with working-class lifestyles and, correspondingly, with the linguistic forms of the working class. This factor may be strong enough to compete successfully against the prestige factor of the linguistic forms associated with more generally prestigious social groups.

## 7.7 Summary

If we ask ourselves why linguistic variables appear to match up with, for example, social class the way that they do an explanation does not seem difficult to come up with. For the most part, people learn their language from members of their own social class and, on the whole, they spend more time talking to those in their own class, or (in a more fine-grained analysis), members of immediately adjacent classes, than they do to people from wildly different social sets. In this sense, the explanation for the relationship between class and linguistic variables is the same as that for the relationship between geographic area and linguistic variables: people usually talk to and learn their linguistic forms from people who live nearby. One obviously tends to sound like those to whom one speaks and from whom one learns to speak. Class, in this sense, is just a measure of 'social space,' in which the working classes are 'more distant' from the upper classes than are the middle classes. 'Social space' regulates (mildly) social interaction, just as geographical space regulates (mildly) social interaction.[4] Variables, like r-lessness, spread from individual to individual in these spaces – geographical and social. After all, it is extremely unlikely that someone who is r-less in New York was born to r-ful parents and just came up with the idea of being r-less on his or her own. An r-less speaker with r-ful parents has almost certainly acquired his r-less dialect by learning it from someone who was already r-less (a schoolmate, for instance). The study of sociolinguistic variables is the study of *diffusion* of linguistic forms or rules throughout the social network.

Why would someone adopt the linguistic forms of another? Obviously, some (if not all) speakers seem to think that they need an additional linguistic system at some point – that their original system was not adequate for all their needs. (Note that neither the decision nor the implementation is conscious. No one suddenly says to him/herself 'Boy, I really need another linguistic system.') What is the motivation for this?

These are difficult questions, though in general there appears to be a single, relatively clear answer: *prestige.* We learn new forms as part of a desire to be more like those who we think are 'cool.' We may adopt various other properties of these prestige individuals, as well: the way they dress, the way they walk, the kinds of things they worry about, etc. Successful adoption of these features requires sufficient contact that we can generate a coherent theory of what system they are using to determine what to wear, how to walk a certain way, and so on. Similarly, in the linguistic dimension, successful learning of the linguistic system of another requires considerable contact. As pointed out above, the bulk of our contact is with those who live in our area and who are members of our own (or closeby) social classes or groups. Every social similarity makes interaction more likely (same gender, same ethnicity, same religion, same class, etc.), and thus opens the door to the learning of a new linguistic system.

Sociolinguistics is thus about the diffusion of linguistic variables through the social network. As such, it can be very revealing as to the precise nature of that network. The origins of the social network are not, of course, the concern of linguists, but what of the origins of linguistic variables? That is, where did the *first* r-less speaker get his or her r-lessness? This question is the proper concern of historical linguistics, to which we now turn.

[4] By 'mildly,' we merely mean not in a deterministic manner – you *can* talk to someone from another place or another class, it's just that on average you are less likely to.

## 7.8 Exercises

1. Which of the following statements is true?

   (a) Isoglosses are lines which show absolute divisions between sets of speakers with certain speech features.

   (b) Isoglosses are lines which indicate natural barriers on maps.

   (c) Isoglosses are lines which represent mid-points in the distribution of some speech feature.

   (d) Isglosses are lines which keep speakers with one speech feature from talking to speakers with another speech feature.

2. Traditional dialectologists focussed on non-urban areas because

   (a) urban areas didn't show any dialect differences.

   (b) speakers in urban areas were too linguistically diverse for there to be a single representation for them.

   (c) speakers in urban areas were too linguistically uniform for it to be an interesting exercise.

   (d) it was difficult to find elderly speakers in urban areas.

3. Which of the following is true?

   (a) The dialect features that you are born with determine what social class you will be in.

   (b) Whatever dialect you are born with, you maintain throughout your life.

   (c) Your dialect features will probably be similar to dialect features of other members of your social class.

   (d) Social variables never correlate with linguistic variables.

4. Which of the following is true?

   (a) Having stigmatized dialect features is a clear indication of intellectual inferiority.

   (b) Having stigmatized dialect features means that you speak only the standard form of the language.

(c) The 'stigmatization' of dialect features is the result of some dialects being better than others from a linguistic point of view.

(d) The 'stigmatization' of dialect features is a social phenomenon, not a linguistic one.

# Chapter 8

# Historical Linguistics

## Contents

**8.1 Language Change** . . . . . . . . . . . . . . . . . . . **164**
8.1.1 Phonological Change . . . . . . . . . . . . . . . . . . 167
8.1.2 Morphological Change . . . . . . . . . . . . . . . . . 169
**8.2 Comparative Reconstruction** . . . . . . . . . . . . **171**
**8.3 Summary** . . . . . . . . . . . . . . . . . . . . . . . . **189**
**8.4 Exercises** . . . . . . . . . . . . . . . . . . . . . . . . **189**

## 8.1 Language Change

Where do the linguistic variables, which diffuse through social networks, come from? Popular media and old fogies will tell you that linguistic change is the result of the decay, laziness, and stupidity of the younger generation. They've said that throughout all of recorded human history (and no doubt earlier than that). While we would be hard pressed to deny those claims for our own generation (let alone younger ones), it seems unlikely that such factors have anything to do with language change. Instead, it appears that change is one of the essential properties of 'language' (in its sociopolitical sense) – we have never observed a linguistic tradition on earth without noting that change (and the variation which results from the diffusion of change) has taken place. Change cannot therefore be something which happens to language as a result of the accidental properties of its speakers (laziness, etc.), which might not be present in a given population or under certain climatic conditions (which tend to be relatively constant for non-nomadic populations). If the universality of change is to be coherently explained, the fact of change must be traced to something inherent in language itself.

It is not difficult to see what feature of language it is that is responsible for universal change. As we pointed out in the previous chapter, the diffusion of sociolinguistic variables is generally the result of learning. You talk in a manner similar to that of the people from whom you learn to talk. The reason for this similarity is to be sought in the mechanism whereby you come to have a linguistic system: while some aspects of that system are innate (as we'll see in the next chapter), much of the details of a particular grammar (e.g., its entire lexicon) must be learned. This learning takes place on the basis of input which you receive as a child – and this input comes from those with whom you interact.

However, as we noted at the beginning of this book, the evidence for the underlying linguistic system that the child receives is seriously flawed all in all. The data is full of performance errors – misspeakings, hesitations, false starts, changes of intended sentence mid-stream, and other lapses – and presented to you not in the sound-proof booth that linguists are fond of doing phonetic research in but in an environment of passing busses, squawking pigeons, annoyingly noisy siblings, people talking with their mouths full, etc. It's a wonder you can extract any coherent system from the awful data you get (a mystery we'll address when we turn to acquisition).

It is hardly surprising, then, that on the basis of the noisy and performance-influenced data which your mother (or whoever) presented you with as a child you failed to converge precisely on the same knowledge state as was in her mind when she was speaking to you. You inevitably ended up with a grammar which was (slightly) different than that of the preceding generation because *language is an imperfectly replicating system.* This imperfect replication is a function of the fact that neither you, nor your mother for that matter, had any form of direct explicit access to the grammatical knowledge in your mother's brain. She could therefore give you no explicit assistant in your challenging learning task (being quite unaware herself of the properties of her grammar – after all, if we all knew what our grammars were like, we'd all get perfect scores on the exams in this class. . . ).

So, the grammar which you construct on the basis of the preceding generations output will differ in subtle ways from the grammars of that generation.

The grammar of your linguistic descendants will also differ from yours in small ways. And the grammar of your descendants' descendants will differ again. Over time, obviously, the differences between generations at some considerable remove from one another will be quite dramatic. For example, many generations ago (e.g., in the fourteenth century) the words corresponding to our [haws] 'house', [maws] 'mouse' and [laws] 'louse' were pronounced [hu:s], [mu:s] and [lu:s]. Small changes from generation to generation since that time have resulted in these words having, for most North Americans, rather different pronunciations.

These examples tell us two important things: first, one type of change is sound change; and second, sound change is systematic and regular. All [u:]'s of Middle English are now pronounced [aw] by us. This regularity allows us to write 'sound change' rules, much as we can write phonological rules (like the rule which aspirates voiceless stops in English) which reside in the grammars of speakers. Examine the data below concerning the development of *h* from Old to Modern English.

a. Old English [hlu:d], Modern English [lawd] 'loud'

b. Old English [hrɪŋ], Modern English [rɪŋ] 'ring'

c. Old English [hnʊtʊ], Modern English [nət] 'nut'

d. Old English [hræfn], Modern English [revən] 'raven'

e. Old English [hnappian], Modern English [næp] 'nap'

f. Old English [hand], Modern English [hænd] 'hand'

g. Old English [he:la], Modern English [hil] 'heel'

h. Old English [hwi:nan], Modern English [wajn] 'whine'

What this data shows (we'll ignore all of the development except those of *h*) is that Old English *h* was sometimes lost and sometimes preserved in Modern English. In fact, it should be clear that the developments can be specified quite precisely: initial *h* was lost before another consonant, but preserved when it preceded a vowel.[1] We can write this as a 'sound change' of the form:

[1] This statement is not true – as you might expect – for all dialects of English. We are using our own dialects as the standard example. For example, some dialects have lost *h*

- h > Ø before consonants

It is important to bear in mind that such rules have a very different status than the rules of theoretical phonology which we discussed earlier. Whereas phonological rules reside in the mind of speakers, regulating the computational processes which govern the generation of phonological forms, sound changes express the relationship between earlier and later grammars (which are assumed to form part of a linguistic tradition). These rules are thus not aspects of any one person's cognitive system.

Change (which is the ultimate source for variation) is not limited to the sound system of a language. We find changes in lexicon (e.g., the loss of 'davenport' as the word for 'couch' or 'sofa' for many North Americans), in morphology (e.g., the replacement of irregular plurals like 'kine' and 'brethren' by regular plurals 'cows' and 'brothers'), syntax, and semantics. A complete survey of these various types of change is beyond the scope of this introductory textbook. We will focus our attention on phonological and morphological change – areas which are relatively well-understood in historical linguistics.

### 8.1.1 Phonological Change

As we saw in the discussion of the development of Old English /h/ above, one of the most striking aspects of phonological change is its *regularity*. Like phonological rules in the grammar, phonological changes apply whenever their conditions are met. Sometimes, these conditions are very general – in such a case, the segment which changes will undergo the change regardless of what phonological environment it is in. Examine the data from the Mafea language of Vanuatu in the table below.[2] The data given in not sufficient to enable you to figure out everything that happened, so we direct your attention to the development of Proto-Oceanic *k.[3]

before all consonants *except* [w] – speakers of these varieties of English still say things like [hwajn] for 'whine'. Other dialects, e.g. Cockney English, have lost *h* in initial position in all words, both before consonants and before vowels.

[2]The IPA symbol ɣ is used to designate a voiced velar fricative. The Mafea sounds designated by p̈ and v̈ are the 'apico-labial' stop and voiced fricative, respectively. 'Apico-labial' designates the very rare articulation in which the tip (apex) of the tongue is placed against (or, in the case of the fricative, very near) the upper lip.

[3]'Proto-Oceanic' is the name given to the *reconstructed* ancestor of most of the languages of Oceania. The method of linguistic reconstruction will be introduced to you later in this chapter. Hypothetical forms reconstructed for proto-languages are marked with an

| Proto-Oceanic | Mafea | English |
|---|---|---|
| **kaɣati* | *ati* | bite |
| **keli* | *eli* | dig |
| **vinsiko* | *v̈isio* | meat |
| **pakiwa* | *p̈aio* | shark |
| **kuɣurita* | *uita* | squid |

As you can see from this data, Proto-Oceanic *k is lost in Mafea in every case – both word-initially and intervocalically.[4] This is an example of an *unconditioned* sound change – the affected segment shows the change in all environments in which it is found. One could write the change as follows:

- *k > Ø

Note the lack of any statement of a conditioning environment for the change.

By way of constrast, examine the data regarding the development of Proto-Oceanic *p in Mafea presented in the table below.

| Proto-Oceanic | Mafea | English |
|---|---|---|
| **pepe* | *p̈ep̈e* | butterfly |
| **apu* | *apu* | fire |
| **puto* | *puto* | navel |
| **ponse* | *poso* | paddle |
| **piri* | *p̈iri* | seed |
| **pakiwa* | *p̈aio* | shark |

We see that Proto-Oceanic *p has two distinct reflexes in Mafea. Since sound change is regular, these two developments must be *conditioned* in some way. An examination of the data above reveals that the conditioning factor is the roundness of the following vowel: Proto-Oceanic *p becomes Mafea p̈ when the following vowel is *not* round, remaining p when the following vowel is round. We can write a formal statement of the change to account for these developments as follows:

- *p > p̈ / __V[-round]

asterisk.

[4]Proto-Oceanic did not have any word-final consonants, so there is no evidence regarding Proto-Oceanic *k in this position.

Note that no change statement need be posited to get Proto-Oceanic *p to come out p in Mafea since there's no change involved in this development, obviously.

In general, the types of events we see in examples of sound change are very much like the types of processes we see in phonological rules. For example, assimilatory changes are very common both as synchronic phonological rules (e.g., English vowel nasalization before nasal vowels) and as diachronic phonological events, as can be seen from the data in the table below.

| Proto-Romance | Portuguese | English |
|---|---|---|
| **kantare* | *kãtar* | to sing |
| **lɔŋga* | *lɔ̃ga* | long |
| **infɛrnu* | *ĩfernu* | hell |
| **dɛnte* | *dɛ̃te* | tooth |

In Portuguese, vowels have become nasalized before nasals (and nasals after vowels are subsequently lost). The first step in this change is assimilatory, just like the English synchronic vowel nasalization rule.

### 8.1.2 Morphological Change

Interestingly, it is not the case that whenever the sounds associated with a particular meaning change, we call it 'sound change.' For example, in earlier forms of English, the plural of 'cow' was 'kine,' the only plural of 'brother' was 'brethren,' and the only comparative of 'old' was 'elder.' The change of earlier English 'kine' to Modern English 'cows' is clearly *not* sound change (e.g., 'mine' did not change into 'mows', 'line' did not change to 'lows', etc.). It seems clear that the change of 'kine' to 'cows,' or of 'brethren' to 'brothers,' is related in some way to the fact that -z is the regular plural morpheme in English. Remember, from our discussion of the 'wug'-test, how we account for an 'irregular' plural like 'oxen' (as opposed to a regular plural, like 'dogs'). The lexical entries of 'ox' and 'dog' include the following information:

- /dɔg/ DOG, [+N], [pl: ]
- /aks/ OX, [+N], [pl: /aksn̩/]

In these entries we see that 'dog' includes a phonemic specification of the underlying phonological form (/dɔg/), a representation of the meaning (DOG), a syntactic category designation ([+N] for nouns), and an indication that there is no idiosyncratic plural form ([pl: ]). By contrast, for 'ox' we find a phonemic representation (/aks/), a semantic representation (OX), a syntactic category designation ([+N]), and an indication that there is an idiosyncratic plural associated with this lexical item ([pl: /aksn̩/]). In the 'wug' test, since the subject of the experiment is assumed not to have a lexical entry 'wug' at all until the experimentor presents them with the word, it seems certain that the subject does not have any 'idiosyncratic' plural information stored with this word. Having no idiosyncratic information stored with a given noun means, of course, that its lexical entry looks just like that of 'dog' with respect to the plural. Like 'dog', such forms get the 'default' English plural marker /-z/, thus /wəgz/ 'more than one wug' is generated.

Looked at this way, we can envision the change of 'cow:kine' to 'cow:cows' in the following way:

- /kaw/ COW, [+N], [pl: /kajn/] → /kaw/ COW, [+N], [pl: ]

In this scenario, the idiosyncratic plural information stored with the lexeme 'cow' changes from having an entry ([pl: /kajn/]) to lacking one ([pl: ]). Since all lexemes which lack idiosyncratic designation of their plural get the regular /-z/ plural in English, /kawz/ will result.

How did the lexeme come to lack its idiosyncratic plural form? We don't know for certain in this particular case, but several possibilities are readily available. Perhaps the learner responsible for this innovation never saw or heard anything about more than one cow until he or she was thirty years old. Not having heard /kajn/, s/he can hardly have stored it in his/her lexicon! Or perhaps a learner heard /kajn/ but assumed it was a generic word for 'cattle' that had nothing to do with 'cow.' This, too, would lead to him or her having no /kajn/ in the plural designation of 'cow' and thus lead him or her to say 'cows' in this function.

The case of 'brethren' is a little more complicated. Imagine that the original situation was that there was a word 'brother' which meant simply 'male sibling'. It would have had a lexical entry like:

- /brəðr̩/ MALE SIBLING, [+N], [pl: /brɛðrən/]

Imagine further that this word could be used metaphorically to refer to any pair of male associates with a particularly close relationship, such as, for example, male members of a monastic order. Eventually, this metaphorical use became so regular that it was taken to be simply a different word altogether. At this point there were two words 'brother' in English, one referring to a male sibling, one to a male member of a religious order. When the original lexical entry split in two, it had to be determined by the learner which one of the forms the irregular plural was associated with. Obviously, some learner got this 'wrong', assigning the irregular plural 'brethren' only to the religious use of the term, leaving the plural designation for the 'male sibling' use of the term blank. A blank entry under '[pl: ]' means that the regular /-z/ plural will be used. We can designate the result of this change to the lexical entry cited above as follows:

- /brəðr̩/ MALE SIBLING, [+N], [pl: ]
- /brəðr̩/ MALE MEMBER OF A RELIGIOUS ORDER, [+N], [pl: /brɛðrən/]

The vast majority of morphological changes involve 'regularization' of forms of this type. The process is sometimes called 'analogy,' based on the idea that the new plurals are created by a form of reasoning usually designated by that name. The form of reasoning works like this

- /ston/:/stonz/::/kaw/:X

which is usually read 'stone is to stones as cow is to X.' If you solve this 'equation' for X you will get the form 'cows.' Unfortunately, this theory can be taken as implying that learners create lexical items by this form of (non-linguistic) reasoning. We have shown above that there are probably more plausible *linguistic* mechanisms which can account for the phenomenon of morphological change.

## 8.2 Comparative Reconstruction

It is often the case, when we examine data from a set of languages, that they show similarities in the forms of morphemes. There are three possible explanations for such similarities: (1) the similarities may be due to simple *chance* – after all, only so many phonological segments exist in human language, and, by chance alone, we might expect *some* relatively small number of morphemes to have similar segments in them when we compare two languages; (2) the similarities may be due to *borrowing* – languages sometimes borrow words from one another, and when they do so, the words will end up looking similar or even identical to one another; or (3) the similarities may be due to the fact that the languages being examined were, at some earlier historical period, *the same*. This last will occur when differences arose due to separate sets of – the phonological or morphological changes diffused throughout a language community leading to the development of 'separate' languages. This last possibility is known as the *genetic* hypothesis – where 'genetic' refers not to some physiological connection through the genes, but rather to the 'origins' meaning of 'genetic' seen in words like 'genesis.' Languages which show genetic, as opposed to chance or borrowing, similarities are known as 'related' languages. Thus, to claim that English is 'related' to Hindi (as it is) is to claim that English and Hindi have evolved by differentiating changes from a common ancestor.

Historical linguistics provides us with a technique for determining attributes of the common ancestor of a set of related languages: the comparative method. We will demonstrate this technique by examining a set of data from two related Micronesian languages – Marshallese and Kiribati. The data we will be dealing with is presented in the table below. There are a few new IPA symbols here, which we will now explain. As we saw in the discussion of phonemic contrasts in Chapter 3, Marshallese has several distinctively *round* consonants: /$r^w$/, /$ŋ^w$/, /$m^w$/, and /$p^w$/ (the last of these is also found in Kiribati). These symbols have their usual interpretation, plus lip-rounding. Thus, $ŋ^w$ is a voiced velar nasal with lip-rounding, $p^w$ is a voiceless labial stop with lip-rounding, etc. Marshallese also has a voiceless palatalized alveolar stop (/$t^y$/) – palatalization involves moving the central part of the tongue up towards the palate. Finally, a dash (-) at the end of a Marshallese word indicates that the form is *never* found without a suffix attached to it – the segment before the dash is thus never word-final in actual Marshallese words.

| | Marshallese | Kiribati | English |
|---|---|---|---|
| 1 | man | man | animal |
| 2 | metak | maraki | ache, pain |
| 3 | kapit | kapira | anoint |
| 4 | $r^w it^y$ | uti | arise |
| 5 | $met^y$ | mata | awake |
| 6 | $nat^y i$- | nati | child |
| 7 | $t^y$aŋit | taŋira | cry about |
| 8 | $met^y a$- | mata | eye |
| 9 | $t^y$iŋ | tiŋ | fart |
| 10 | $t^y$ema- | tama | father |
| 11 | lima- | nima- | five |
| 12 | $laŋ^w$ | naŋo | fly |
| 13 | $p^w il$ | $p^w uni$ | gum, sap |
| 14 | $r^w eŋ^w$ | oŋo | hear |
| 15 | $nam^w$ | namo | lagoon |
| 16 | teŋa- | raŋa | leg |
| 17 | $nek^w$ | noko | midrib of a coconut frond |
| 18 | likit | nikira | put |
| 19 | $p^w ap^w$ | $p^w apu$ | shark species |
| 20 | kaŋ | kaŋ | sharp |
| 21 | $mat^y ir^w$ | matuu | sleep |
| 22 | metal | maran | smooth |
| 23 | $t^y ir$ | tii | spurt |
| 24 | lewe- | newe | tongue |
| 25 | til | rin | torch |
| 26 | pit | piro | twist sennit |
| 27 | tak | rake | upward |
| 28 | $rat^y$ | ato | whale |

First note that these Marshallese and Kiribati forms show a remarkable degree of similarity – contrast these forms with the forms in the English column, for example, and you will notice that they are much more like one another than either is like the English forms. Moreover, although it would take too much space to demonstrate it conclusively here, this similarity between Marshallese and Kiribati is pervasive – it holds of hundreds and hundreds more vocabulary items than we have presented. This similarity is too pervasive for it to be attributed to borrowing – quite aside from the fact that the Marshallese and Kiribati speakers are separated by hundreds of miles of open ocean (making 'borrowing' words from one another quite a challenge!). In addition, the similarity is too wide-spread for it to be due to chance. This leaves us with a genetic explanation for the similarities – Marshallese and Kiribati must have once been the same language, the current differences between them being due to a set of phonological changes which have taken place in each language since they went their separate ways (much as North American English has come to be different from British English).

Imagine that you were asked to formulate a scientific hypothesis about what this common ancestor of Marshallese and Kiribati (call it 'Proto-Micronesian') looked like. Let us start with the first piece of data in the table above: Marshallese *man* and Kiribati *man* mean 'animal.' What do we believe the Proto-Micronesian word for 'animal' probably was? It seems that hypothesizing that this word was anything but **man* would be pretty silly.

We can, in fact, break down what we just did in coming up with Proto-Micronesian **man* 'animal' into smaller steps. The reasoning that was used to arrive at this theory looks something like this:

| Marshallese | Kiribati | Proto-Micronesian |
|---|---|---|
| m | m | *m |
| a | a | *a |
| n | n | *n |

The relationships represented in this table are known as 'sound correspondences.' The table can be interpreted as follows: an *m* in Marshallese corresponds to an *m* in Kiribati, and this correspondence is to be explained as derived from a Proto-Oceanic **m*. We can also extract from this a simple principle:

- Principle I: If all the languages being compared have the same segment in a given correspondence, reconstruct that segment for the ancestral proto-language.

Note that this merely captures the logic which you yourself used when you were asked what you would hypothesize the Proto-Oceanic word for 'animal' was. It represents a binding principle: if all of the languages being compared have a given segment in a correspondence set, you are not allowed, in reconstructing the proto-language, to posit any other segment.

If we examine more of the data, we will discover more correspondence sets where we find the same segment in both Marshallese and Kiribati. Look at the table below.

| Marshallese | Kiribati | Proto-Micronesian | example numbers |
|---|---|---|---|
| a | a | *a | 1, 2, 3, 6, 7, 8, 10, 11, 12, 15, 16, 19, etc. |
| e | e | *e | 24 (twice) |
| i | i | *i | 3, 6, 7, 9, 11, 18 (twice), 23, 25, 26 |
| m | m | *m | 1, 2, 5 (twice), 8, 10, 11, 21, 22 |
| n | n | *n | 1, 6, 15, 17 |
| ŋ | ŋ | *ŋ | 7, 9, 16, 20 |
| p | p | *p | 3, 26 |
| $p^w$ | $p^w$ | *$p^w$ | 13, 19 |
| k | k | *k | 2, 3, 18, 20 (twice), 27 |
| w | w | *w | 24 |

The information in the table can be read as follows. If you examine, e.g., the Marshallese ŋ in example 7 (***$t^y$aŋit*** 'cry about'), you will find in the corresponding word in Kiribati (**taŋira**) an ŋ as well, and that given the principle stated above, Proto-Micronesian should have an ŋ at this position in this word as well. The correspondences in this table are found to recur throughout the vocabularies of these two languages, though we cannot, for reasons of space, demonstrate this here.

If all of the correspondences were identical in this way, the vocabularies of the two languages would be the same. This is obviously not true here, so we need to examine the cases in which the correspondences do *not* involve identity in form in Marshallese and Kiribati. The process used to do this is described below.

As can be seen in the table above, we find a richly attested correspondence of identity between Marshallese *i* and Kiribati *i* (reflecting, given the principle stated above, Proto-Micronesian **i*), we also find the correspondence in the table below.

| Marshallese | Kiribati | Proto-Micronesian | example numbers |
|---|---|---|---|
| Ø | i | ??? | 2, 4, 13, 23 |

There are two possible explanations for why Kiribati *i* sometimes corresponds to Marshallese *i* and sometimes corresponds to Marshallese Ø, correlating with the contrast we drew above between *conditioned* and *unconditioned* sound change. The first possibility is that this correspondence also represents Proto-Micronesian **i*, which was lost in Marshallese in some environments (giving the Ø:i:*i correspondence) and preserved in others (giving the i:i:*i set). The second possibility is that the correspondence set in the table above represents some segment other than *i (e.g., *ɪ), which unconditionally was lost in Marshallese and unconditionally became *i* in Kiribati. The general principle we use in a case like this is the following:

- Principle II: Minimize the number of segments reconstructed for the proto-language.

This principle implies that *if* we can derive the Marshallse Ø = Kiribati *i* correspondence by the use of a conditioned sound change, we should do so. To determine whether we can do so, we need to look at the phonological environments in which we find, on the one hand, the Marshallese *i* = Kiribati *i* correspondence and, on the other hand, the Marshallese Ø = Kiribati *i* correspondence. The relevant data is the following:[5]

- Marshallese *i* = Kiribati *i*: p_t (3), tʸ_- (6), ŋ_t (7), tʸ_ŋ (9), l_m (11), l_k (18), k_t (18), tʸ_r (23), t_l (25), p_t (26)
- Marshallese Ø = Kiribati *i*: k_# (2), tʸ_# (4), l_# (13), r_# (23)

Notice that all of the cases of the Marshallese Ø = Kiribati *i* correspondence set involve an *i* that is word-final in Kiribati and that *none* of the cases of Marshallese *i* = Kiribati *i* involve word-final *i*.[6] This means that we can easily write a rule which will derive this correspondence:

- *i > Ø / __# in Marshallese

What this rule tells us is that Proto-Micronesian **i* is lost in word-final position in Marshallese. When not in word-final position, it is not lost. Therefore, Proto-Micronesian **i* has two reflexes in Marshallese: Ø in word-final position, *i* elsewhere. We can see then that conditioned sound changes will give two reflexes of the proto segment in a given language. The table for this correspondence should have the form below:

[5] We will use the Marshallese forms to describe the environment, though this doesn't matter much in the present context.

[6] Remember that the dash at the end of Marshallese *natʸi-* 'child' means that this morpheme is always followed by a suffix – thus its *i* is never word-final.

| Marshallese | Kiribati | Proto-Micronesian | Rule |
|---|---|---|---|
| i | i | *i | none needed |
| Ø | i | *i | *i > Ø / __# in Marshallese |

By contrast, unconditioned change will give only one reflex of the proto segment. To see an example of that, let us examine another pair of correspondence sets found in the data.

| Marshallese | Kiribati | Proto-Micronesian | example numbers |
|---|---|---|---|
| n | n | *n | 1, 6, 15, 17 |
| l | n | ??? | 11, 12, 13, 18, 22, 24, 25 |

The first correspondence set *must* be Proto-Micronesian *n by Principle I, but what of the second? Again, there are two possibilities: either the second correspondence set also reflects Proto-Micronesian *n, the Marshallese *l* being due to a conditioned sound change (*n > l in some environment or another), or we are dealing with a Proto-Micronesian segment which was different from *n. If we examine the environments in which these correspondence sets are found, we see that in the word for 'child' (example 6) the n=n correspondence set is found in the environment #__a (i.e., word-initially before *a*). However, in the word for 'fly' (example 12), we find that the correspondence set l=n is found in precisely that same environment. This means that there is no way we can write a rule which would change Proto-Micronesian *n to Marshallese *l* in 'fly' without it also changing the Proto-Micronesian *n of 'child' to Marshallese *l* – which we do not want it to do (since the Marshallese word for 'child' has an *n*). Therefore we have to posit a segment other than **n* in Proto-Micronesian for the second correspondence set in the table above. But what segment? Well, it seems logical that we should reconstruct *l: after all, Marshallese has *l* and Kiribati has *n*, and we know the segment wasn't **n* (by the reasoning we just went through). So if the most logical choices are either it was **l* (since it is *l* in Marshallese) or it was **n* (since it is *n* in Kiribati), and we know it wasn't **n*, it must have been **l*! If it was **l*, then of course we need it to change to *n* in Kiribati (since that how it shows up in the data), i.e., we need the rule *l > n in Kiribati (unconditionally). Since the rule has no conditions, there is only one reflex of Proto-Micronesian **l* in Kiribati. Thus the table above can be completed as:

| Marshallese | Kiribati | Proto-Micronesian | Rule |
|---|---|---|---|
| n | n | *n | none needed |
| l | n | *l | *l > n in Kiribati |

The basic question which arises when we see correspondence sets which do not involve identity of forms in languages being compared is whether the correspondence set points to a conditioned development of a proto-segment we need anyway (and thus Principle II tells us we must reconstruct the same segment here) or whether it points to an unconditioned development of some other proto-segment. We determine this by an examination of the environments within which the correspondence set is found: if these environments are in complementary distribution with the environments in which some other similar correspondence set is found, we are probably looking at a conditioned sound change (as in the Marshallese *i-loss example above). If, by contrast, the environment overlaps with that of the other similar correspondence set, we are looking at two distinct segments in the proto-language (as in the Kiribati *l > n case just discussed).

Let us turn to some other readily resolved aspects of the Micronesian data. We find several correspondence sets involving Ø in one of the languages. One of these correspondence sets is given in the table below.

| Marshallese | Kiribati | Proto-Micronesian | example numbers |
|---|---|---|---|
| r | Ø | ??? | 23, 28 |

Following the logic briefly mentioned above, we have the choice here between reconstructing Proto-Micronesian *r (as in Marshallese) or Proto-Micronesian *Ø (as in Kiribati). However, if we were to reconstruct *Ø, we would have to state the conditions under which Marshallese speakers inserted *r* when there was none in the proto-language. Again, a general principle is helpful here:

- Principle III: Deletion is highly preferred over insertion.

This principle is based on the observed fact that languages delete segments in the course of their history far more often than they insert segments. Given any correspondence set involving Ø, therefore, we should prefer one of the non-zero candidates to positing a Ø in the proto-language. That is, the table involving Kiribati Ø above should be:

| Marshallese | Kiribati | Proto-Micronesian | Rule |
|---|---|---|---|
| r | Ø | *r | *r > Ø in Kiribati |

This principle also helps us deal with several other correspondence sets present in the data. Examine the table below.

| Marshallese | Kiribati | Proto-Micronesian | example numbers |
|---|---|---|---|
| Ø | o | ??? | 12, 14, 15, 17, 26, 28 |
| Ø | e | ??? | 27 |
| Ø | u | ??? | 19, 21 |
| Ø | a | ??? | 3, 5, 7, 18 |

Given Principle III, these examples are most likely to involve deletion in Marshallese, rather than insertion in Kiribati. Indeed, we discussed the Marshallese Ø = Kiribati *i* correspondence set above and concluded that it involved conditioned loss of Proto-Micronesian **i* in word-final position. In every one of the cases above, the correspondence set involves what is in Kiribati a word-final vowel, showing up as Ø in Marshallese. Thus, a very general rule appears to be in order, one which deletes all Proto-Micronesian word-final vowels in Marshallese. This general process is known as 'final vowel deletion.'

- V > Ø / __# in Marshallese

We can now combine the data from the table above with our earlier observations concerning the Marshallese Ø = Kiribati *i* correspondence set in the following table:

| Marshallese | Kiribati | Proto-Micronesian | Rule |
|---|---|---|---|
| Ø | a | *a | final vowel deletion in Marshallese |
| Ø | e | *e | final vowel deletion in Marshallese |
| Ø | i | *i | final vowel deletion in Marshallese |
| Ø | o | *o | final vowel deletion in Marshallese |
| Ø | u | *u | final vowel deletion in Marshallese |

We also find in the data several cases in which Marshallese has a round consonant and Kiribati has a corresponding plain (i.e., non-round) consonant. These examples are given in the table below.

| Marshallese | Kiribati | Proto-Micronesian | example numbers |
|---|---|---|---|
| ŋʷ | ŋ | ??? | 12, 14 |
| mʷ | m | ??? | 15 |
| kʷ | k | ??? | 17 |
| pʷ | p | ??? | 19 |

Taking the ŋʷ case first, we can see that either this correspondence set represents a conditioned development of Proto-Micronesian *ŋ (which was reconstructed for examples in which both languages show *ŋ* – 7, 9, 16, 20), or we must posit an additional segment (*ŋʷ) for the proto-language. Principle II says that we should not expand the set of segments in the proto-language unless we have to, so we need to determine whether Marshallese ŋʷ could be conditioned development of Proto-Micronesian *ŋ. The Kiribati form of the words in question is *naŋo* 'fly' and *oŋo* 'hear.' We can see that both of these *ŋ*'s occur before the round vowel *o* and show up as round in Marshallese (*laŋʷ* 'fly' and *rʷeŋʷ* 'hear'). By contrast, none of the cases of the correspondence set Marshallese ŋ = Kiribati ŋ involve an ŋ which is before a round vowel. We can therefore derive the Marshallese *ŋʷ*'s with a rule such as:

- *ŋ > ŋʷ / __V[+round] in Marshallese

Similarly, if we examine the case of Marshallese *mʷ* = Kiribati *m* we see that the Kiribati form of the word in question is *namo*, with a round vowel following the *m*. Again, none of the many cases of the correspondence set Marshallese *m* = Kiribati *m* involve an *m* which is before a round vowel. We can thus posit the rule:

- *m > mʷ / __V[+round] in Marshallese

You can easily confirm that the very same generalization is true for the remaining examples of round consonants in Marshallese corresponding to non-round consonants in Kiribati. In every case, the round consonants of Marshallese which are in these correspondence sets are found before round vowels. In fact, there appears to be a general rule at work here: Proto-Micronesian **k*, **p*, **m* and **ŋ* become round before round vowels.

There are only a few correspondence sets left to consider. Two of them are given in the table below.

| Marshallese | Kiribati | Proto-Micronesian | example numbers |
|---|---|---|---|
| t | r | ??? | 2, 3, 7, 1, 18, 22, 25, 26, 27 |
| $t^y$ | t | ??? | 4, 5, 6, 7, 8, 9, 10, 21, 23, 28 |

These are clearly similar correspondence sets, so we must first confront the question of whether or not they could represent conditioned developments of a single original segment (e.g., *t). As pointed out above, this involves looking at the examples which contain the correspondence set to see if there is complementary distribution in the environments in question. Look at the forms in the table below.

| | Marshallese | Kiribati | English | relevant correspondence |
|---|---|---|---|---|
| 25 | til | rin | torch | t=r |
| 23 | $t^y$ir | tii | spurt | $t^y$=t |
| 27 | tak | rake | upward | t=r |
| 7 | $t^y$aŋit | taŋira | cry about | $t^y$=t |

As is clear from that data (repeated from the large data table at the beginning of this section), it would be impossible to collapse these forms under a single proto-segment. Imagine that we claimed that both of the correspondence sets reflect a conditioned development of Proto-Micronesian *t. To get the forms in 25 and 27 in the table above, we would need Proto-Micronesian *t to become Kiribati *r* in the following environments: word-initially before *i* (in 25) and word-initially before *a* (in 27). We would also need for Proto-Micronesian *t to remain *t* in these same environments in Marshallese (as shown by the Marshallese forms in 25 and 27, which have *t*). However, to get the Marshallese forms in 23 and 7 we would need Proto-Micronesian *t to become Marshallese $t^y$ in the following environments: word-initially before *i* (in 23) and word-initially before *a* (in 7). However, as we just saw, to account for the t:r correspondence set, Proto-Micronesian *t would need to remain *t* in Marshallese in precisely these same environments. Since Proto-Micronesian *t cannot both remain *t* and become $t^y$ in the same environment, the correspondence sets in the table above cannot be derived from a single proto-segment.

So we must be dealing with two distinct Proto-Micronesian segments in the Marshallese *t* = Kiribati *r* and Marshallese $t^y$ = Kiribati *t* correspondence sets. But which two segments? For the correspondence set Marshallese *t* = Kiribati *r* there are two logically possible choices: **t* or **r*. However, as we saw above in our discussion of the Marshallese *r* = Kiribati Ø correspondence set, for which we reconstructed Proto-Micronesian **r*, Proto-Micronesian **r* is lost in Kiribati. The proto-language segment reflected in the *t*=*r* correspondence set therefore cannot be Proto-Micronesian **r*. Since the only logical choices were either **t* or **r*, and since it cannot have been **r*, it must have been **t*. This means of course that Proto-Micronesian **t* became *r* in Kiribati, as in the table below.

| Marshallese | Kiribati | Proto-Micronesian | Rule |
|---|---|---|---|
| t | r | *t | *t > r in Kiribati |

Turning next to the Marshallese $t^y$ = Kiribati *t* correspondence set we again note that there are two logical possibilities: Proto-Micronesian $*t^y$ or Proto-Micronesian **t*. However, we just reconstructed Proto-Micronesian **t* for the correspondence set Marshallese *t* = Kiribati *r*, and we demonstrated above that the correspondence set we are now discussing must reflect a *different* segment than that of the Marshallese *t* = Kiribati *r* set. The Marshallese $t^y$ = Kiribati *t* set thus cannot represent Proto-Micronesian **t* (which we've just seen becomes *r*, not *t*, in Kiribati). Since it must logically either represent **t* or $*t^y$, and it cannot represent **t*, by necessity it must reflect Proto-Micronesian $*t^y$. This means that Proto-Micronesian $*t^y$ > Kiribati *t*. The relevant correspondences are shown in the table below.

| Marshallese | Kiribati | Proto-Micronesian | Rule |
|---|---|---|---|
| r | Ø | *r | *r > Ø in Kiribati |
| t | r | *t | *t > r in Kiribati |
| $t^y$ | t | $*t^y$ | $*t^y$ > t in Kiribati |

We'll take a little breather at this point and see how much of the data in the original table we have now figured out. The forms which are accounted for by the reconstructions we've made so far are given in the table below. (We'll use the following abbreviations to make the table fit on the page better: K=Kiribati, M=Marshallese, PMC=Proto-Micronesian, FVD=final vowel deletion.)

| Marshallese | Kiribati | PMC | English | rules needed |
|---|---|---|---|---|
| man (1) | man | *man | animal | none |
| kapit (3) | kapira | *kapita | anoint | *t>r in K, FVD in M |
| met$^{y}$ (5) | mata | *mat$^{y}$a | awake | *t$^{y}$>t in K, FVD in M |
| nat$^{y}$i- (6) | nati | *nat$^{y}$i | child | *t$^{y}$>t in K |
| t$^{y}$aŋit (7) | taŋira | *t$^{y}$aŋita | cry about | *t$^{y}$>t & *t>r in K, FVD in M |
| t$^{y}$iŋ (9) | tiŋ | *t$^{y}$iŋ | fart | *t$^{y}$>t in K |
| lima- (11) | nima- | *lima | five | *l>n in K |
| laŋ$^{w}$ (12) | naŋo | *laŋo | fly | *l>n in K, *ŋ-rounding & FVD in M |
| nam$^{w}$ (15) | namo | *namo | lagoon | *m-rounding & FVD in M |
| likit (18) | nikira | *likita | put | *l>n & *t>r in K, FVD in M |
| p$^{w}$ap$^{w}$ (19) | p$^{w}$apu | *p$^{w}$apu | shark sp. | *p-rounding & FVD in M |
| kaŋ (20) | kaŋ | *kaŋ | sharp | none |
| t$^{y}$ir (23) | tii | *t$^{y}$iri | spurt | *t$^{y}$>t & r-loss in K, FVD in M |
| lewe- (24) | newe | *lewe | tongue | *l>n in K |
| til (25) | rin | *til | torch | *t>r & *l>n in K |
| pit (26) | piro | *pito | twist sennit | *t>r in K, FVD in M |
| tak (27) | rake | *take | upward | *t>r in K, FVD in M |
| rat$^{y}$ (28) | ato | *rat$^{y}$o | whale | *r>Ø & *t$^{y}$>t in K, FVD in M |

The remaining problematic forms, which we have not yet accounted for, are repeated in the table below so that you won't have to keep flipping the pages. In addition, I have indicated what we know so far about the Proto-Micronesian forms of these words – designating what we don't yet know by a blank underline mark.

| | Marshallese | Kiribati | PMC | English |
|---|---|---|---|---|
| 2 | metak | maraki | *m_taki | ache, pain |
| 4 | r$^{w}$it$^{y}$ | uti | *__ti | arise |
| 8 | met$^{y}$a- | mata | *m_t$^{y}$a | eye |
| 10 | t$^{y}$ema- | tama | *t$^{y}$_ma | father |
| 13 | p$^{w}$il | p$^{w}$uni | *p$^{w}$_li | gum, sap |
| 14 | r$^{w}$eŋ$^{w}$ | oŋo | *__ŋo | hear |
| 16 | teŋa- | raŋa | *t_ŋa | leg |
| 17 | nek$^{w}$ | noko | *n_ko | midrib of a coconut frond |
| 21 | mat$^{y}$ir$^{w}$ | matuu | *mat$^{y}$_ru | sleep |
| 22 | metal | maran | *m_tal | smooth |

In this data we see several examples (numbers 2, 8, 10, 16, and 22) of a correspondence set in which Marshallese *e* = Kiribati *a*. Remember that based on cases of Marshallese *a* = Kiribati *a* we have already reconstructed a Proto-Micronesian **a*, and based on cases of Marshallese *e* = Kiribati *e* we have already reconstructed a Proto-Micronesian **e*. Thus the two logical choices offered by the *e*=*a* correspondence set (**e* and **a*) would each correspond to a prior existing segment in the proto-language. This makes it likely that this correspondence set is due to a conditioned change of either Proto-Micronesian **a* or Proto-Micronesian **e*. To determine which of these is correct, we must examine the conditioning involved.

Here a special consideration comes into play. When examining environments up to now we have limited ourselves to the strictly local environment, in keeping with what we pointed out earlier concerning phonological rules. However, as noted in that discussion, there is a well-attested exception to this strict locality of phonological processes: vowels frequently interact with one another across intervening consonants (i.e., as if the consonants in question were not there). It is just such a factor which is relevant to the problem at hand, for if we examine all of the cases involving the Marshallese *e* = Kiribati *a* correspondence set, we see that in every case the vowel in question was followed, in the next syllable, by a Proto-Micronesian **a*. The examples include *m_t**a**ki 'ache, pain' (2), *m_t$^y$**a** 'eye' (8), *t$^y$_m**a** 'father' (10), *t_ŋ**a** 'leg' (16), and *m_t**a**l 'smooth' (22).

Now, if we were to assume that the Marshallese *e* = Kiribati *a* correspondence set involves Proto-Micronesian **a*, we would need a rule like:

- *a > e / __Ca (where C is any consonant) in Marshallese

whereas if we were to assume that this correspondence set represents a Proto-Micronesian **e*, we would need a rule like

- *e > a / __Ca (where C is any consonant) in Marshallese

How can pick between these two rules, and thus determine whether we are dealing with a Proto-Micronesian **a* or a Proto-Micronesian **e*? Because assimilation processes are far more frequent than dissimilation processes in the historical phonology of human languages, we have the following principle to help us:

- Principle IV: If you must choose between positing an assimilation or a dissimilation, select the assimilation.

Since in the first proposed rule above a Proto-Micronesian **a* becomes *less like* the triggering **a* of the following syllable this is a dissimilation rule. By contrast, in the second proposed rule above, a Proto-Micronesian **e* becomes *more like*, i.e. assimilates to, the **a* of the next syllable. The second rule is therefore to be preferred by Principle IV, and we must reconstruct Proto-Micronesian **e* for these forms. This gives the following result for some of the problematic forms in the table above.

| | Marshallese | Kiribati | PMC | English |
|---|---|---|---|---|
| 2 | metak | maraki | *metaki | ache, pain |
| 8 | met$^y$a- | mata | *met$^y$a | eye |
| 10 | t$^y$ema- | tama | *t$^y$ema | father |
| 16 | teŋa- | raŋa | *teŋa | leg |
| 22 | metal | maran | *metal | smooth |

The next set of problematic forms we must deal with are those involving Marshallese **r$^w$* (numbers 4, 14, and 21 in the table above). At first glance, this issue seems trivial: we are dealing with a correspondence set Marshallese r$^w$ = Kiribati Ø, and we have a principle that says it is better to delete than to insert, so it seems to follow that we should reconstruct **r$^w$* here. However, we have already seen that Marshallese has a rule of consonant rounding, whereby certain consonants are rounded before round vowels. So another possibility concerning Marshallese r$^w$ arises: perhaps it reflects Proto-Micronesian **r* in Marshallese consonant-rounding environments (i.e., before round vowels). We can now see how this problem relates to the sole remaining difficulty in the data we have been considering: there are several cases in which Marshallese non-round vowels correspond to Kiribati round vowels. These include the correspondence sets Marshallese *i* = Kiribati *u* (in examples 4, 13, and 21) and Marshallese *e* = Kiribati *o* (in examples 14 and 17). The first thing to note about this vowel problem is that if we were to reconstruct the non-round vowels for these correspondence sets it would turn out that the only round vowels in Proto-Micronesian would be those which occured in word-final position (as in, for example, Proto-Micronesian **laŋo* 'fly', **namo* 'lagoon', and **p$^w$apu* 'shark species'). This would be very strange, since we don't find human languages with such a funny distribution of round vowels anywhere on earth today.

Secondly, we know in the case of the word for 'sleep' that by our reconstruction it ended in a Proto-Micronesian **u* and Marshallese has a round

$r^w$ in this word – so it looks like the roundness of Marshallese $r^w$ may in fact be due to following round vowels of Proto-Micronesian. But if this analysis were to be correct, then the vowels following the $r^w$ in Marshallese $r^w it^y$ 'arise' and $r^w e\eta^w$ 'hear' must have been round. These vowels are precisely members of the Marshallese $i$ = Kiribati $u$ and Marshallese $e$ = Kiribati $o$ correspondence sets. So, if we reconstruct these vowels as round (i.e., with their Kiribati values), we can get rid of $*r^w$ as segment in the proto-language (in keeping with Principle II). Note that, by contrast, if we reconstruct these vowels as non-round (in keeping with their Marshallese values), we must have a segment $*r^w$ in the proto-language, and we also must have the segments $*u$ and $*o$, because of their occurrence in word-final positions in the words for 'fly', 'lagoon', 'shark species', 'twist sennit' and 'whale.'

Thus only by selecting the Kiribati values for the correspondence sets Marshallese $i$ = Kiribati $u$ and Marshallese $e$ = Kiribati $o$ can we obey Principle II and limit the set of segments we establish for the proto-language. The ultimate resolution of the 'problematic' forms is as given in the table below.

| Marshallese | Kiribati | PMC | English | rules needed |
|---|---|---|---|---|
| metak (2) | maraki | *metaki | ache, pain | *t>r & *e>a / __Ca in K, FVD in M |
| $r^w it^y$ (4) | uti | *ruti | arise | $*t^y$>t & r-loss in K, C-rounding, *u>i & FVD in M |
| $met^y a$- (8) | mata | $*met^y a$ | eye | $*t^y$>t & *e>a / __Ca in K |
| $t^y ema$- (10) | tama | $*t^y ema$ | father | $*t^y$>t & *e>a / __Ca in K |
| $p^w il$ (13) | $p^w uni$ | $*p^w uli$ | gum, sap | *l>n in K, *u>i & FVD in M |
| $r^w e\eta^w$ (14) | oŋo | *roŋo | hear | *r>Ø in K, C-rounding, *o>e & FVD in M |
| teŋa- (16) | raŋa | *teŋa | leg | *t>r & *e>a / __Ca in K |
| $nek^w$ (17) | noko | *noko | midrib | *o>e, C-rounding & FVD in M |
| $mat^y ir^w$ (21) | matuu | $*mat^y uru$ | sleep | $*t^y$>t & r-loss in K, C-rounding, *u>i & FVD in M |
| metal (22) | maran | *metal | smooth | *l>n, *t>r & *e>a / __Ca in K |

The sound changes we have posited for Kiribati are the following:

- *e > a / __Ca (where C is any consonant)
- $*t^y$ > t

- *t > r
- *r > Ø
- *l > n

The sound changes we have posited for Marshallese are the following:

- V > Ø / __# (final vowel deletion)
- round vowels > [-round] (i.e., *u > i and *o > e)
- *p, *k, *m, *ŋ, and *r > [+round] before round vowels (consonant rounding)

We have not paid any attention, up to this point, to what *order* the sound changes must have applied in. That this is relevant can be seen from a simple example. We have reconstructed the word for 'arise' as Proto-Micronesian **rut$^y$i*. Imagine that we were to have the sound changes of Kiribati take place in the order given above to this form. We start with the proto-form **rut$^y$i*. The first rule is about **e* and there aren't any of these in the form, so the rule is irrelevant. The next rule says that **t$^y$* > t. This rule will change our form into **ruti*.[7] The next rule says that **t* becomes *r* – note that our form now has a **t*. It would thus become **ruri*. The next rule triggers the loss of all *r*'s; our form would therefore become *ui*. The final rule is irrelevant. The rules applied in this order predict that the Kiribati word for 'arise' should be *ui*, which it most definitely is not.

Similarly, if we start with Proto-Micronesian **rut$^y$i* and apply the Marshallese changes in the order stated, we will get a chain of events like the following. The first change will delete the final vowel of our proto-form, giving **rut$^y$*. The next change will unround the round vowels, giving **rit$^y$*. The third rule, consonant rounding, will not apply because the vowel following the *r* in our form is, at this point, no longer round. Thus the Marshallese form is predicted, incorrectly, to be *rit$^y$*, whereas it is in fact ***r$^w$it$^y$***.

These two examples reveal the importance of not merely discovering what changes have taken place in the phonological history of a given language, but

[7]The form continues to get an asterisk, because the asterisk designates that the form is reconstructed, rather than found in an actually attested language.

also of identifying what order these changes must have taken place in. Sometimes determining the ordering is not possible (e.g., we cannot determine at what point Proto-Micronesian **l* became Kiribati *n*), but when it is possible, we must state the ordering explicitly, to prevent the derivation of the wrong form.

In general, we must prevent rules from inappropriately feeding one another – feeding is a process whereby a rule creates a segment which fits the environment for a later rule – or inappropriately bleeding one another – bleeding is a process whereby a rule destroys the environment for a later rule. For example, in the Kiribati case sketched above, the problem arises because the change of **t*$^y$ to *t* inappropriately feeds the later change of **t* to *r* (which itself inappropriately feeds the r-loss rule). If, instead of the order above, we have the order

- *r > Ø
- *t > r
- *t$^y$ > t

there will be no inappropriate feeding relationships between the rules because **r* will have gotten lost **before** **t* becomes *r*, and **t* will have become *r* **before** **t*$^y$ becomes *t*.

In the Marshallese case we are dealing with an instance of inappropriate bleeding of one rule by another. If the rule which unrounds round vowels applies before the consonant rounding rule, there will be no round vowels before which the consonants can become round! Since we know the consonants do in fact become round, the order of these two rules must be (1) consonant rounding, then (2) vowel unrounding. What about final vowel deletion? Notice that if we apply the rule of final vowel deletion before the rule of consonant rounding, then a form like Proto-Micronesian **namo* will come out as Marshallese *nam*:

- Proto-Micronesian *namo
- final vowel deletion produces *nam
- consonant rounding has nothing to apply to!

Since this is not the correct outcome (consonant rounding has been inappropriately bled by final vowel deletion), we know that final vowel deletion must **follow** consonant rounding in Marshallese. No order can be determined between final vowel deletion and vowel unrounding.

## 8.3 Summary

The very complicated example which we have just worked through an important phenomenon that we mentioned earier in the chapter: sound change is regular. The similarities between Marshallese and Kiribati can be explained by starting with a well-defined set of Proto-Micronesian forms and applying to these an ordered set of regular changes, different for each of the two languages. It is the regularity of change which makes it possible for us to demonstrate that languages are related and thus that they fall into groups of relatedness, called 'language families.' Some of the properties of the languages ancestral to the set of languages within a family can be determined by the use of the comparative method, which thereby reveals aspects of the history of human linguistic systems which we could not discover in any other way.

The evidence uncovered by linguistic reconstruction can be helpful in determining where a given linguistic group originates. For example, imagine that from a set of languages you determine that their ancestor had a particular word for 'snow,' 'hail,' 'ice,' 'polar bear,' and 'ice-fishing hut.' It is most unlikely that the linguistic ancestor of these languages was spoken on a South Pacific island, for example. Any theory which proposes that these people have migrated from some temperate climate is not likely to be correct. Linguistic reconstruction can reveal aspects of the culture and living environment of the ancestral speakers which could not otherwise be determined.

## 8.4 Exercises

Comparative Polynesian

A. Examine the data below and on the basis of that data reconstruct the common ancestor of Hawaiian, Maori, Samoan and Tongan (this ancestor is called Proto-Polynesian).

B. List the changes required to get from your Proto-Polynesian forms to the attested forms in the languages given below, noting any necessary ordering.

C. What does the set of words which we can reconstruct tell you about the culture and living environment of the Proto-Polynesians?

| Hawaiian | Maori | Samoan | Tongan | English |
|---|---|---|---|---|
| manu | manu | manu | manu | animal, bird |
| niu | niu | niu | niu | coconut |
| inu | inu | inu | inu | drink |
| pua | pua | pua | pua | flower |
| mimi | mimi | mimi | mimi | urinate |
| peʔa | peka | peʔa | peka | bat |
| muli | muri | muli | mui | behind |
| ani | aŋi | aŋi | aŋi | blow |
| kani | taŋi | taŋi | ŋi | cry |
| au | au | au | ʔau | current |
| kuna | tuna | tuna | tuna | eel species |
| walu | waru | walu | walu | eight |
| maka | mata | mata | mata | eye |
| iʔa | ika | iʔa | ika | fish |
| ʔai | kai | ʔai | kai | food, eat |
| kae | tae | tae | taʔe | excrement |
| lono | roŋo | loŋo | oŋo | hear |
| lau | rau | lau | lau | leaf |
| ake | ate | ate | ʔate | liver |
| ʔuku | kutu | ʔutu | kutu | louse |
| limu | rimu | limu | limu | moss, seaweed |
| nuku | ŋutu | ŋutu | ŋutu | mouth, beak |
| piko | pito | pito | pito | navel |
| ihu | ihu | isu | ihu | nose |
| ama | ama | ama | hama | outrigger |
| ʔiako | kiato | ʔiato | kiato | outrigger boom |
| umu | umu | umu | ʔumu | oven, earthen |
| ua | ua | ua | ʔuha | rain |
| ala | ara | ala | hala | road |
| kai | tai | tai | tahi | sea |
| walu | waru | walu | wau | scratch |
| ʔili | kiri | ʔili | kili | skin |
| kapu | tapu | tapu | tapu | taboo |
| ako | ato | ato | ʔato | thatch, roof |
| lua | rua | lua | ua | two |
| lua | rua | lua | lua | vomit |
| nalu | ŋaru | ŋalu | ŋalu | wave |
| lena | reŋa | leŋa | eŋa | yellow, turmeric |

| Hawaiian | Maori | Samoan | Tongan | English |
|---|---|---|---|---|
| ahi | ahi | afi | afi | fire |
| haa | $h^w$aa | faa | faa | four |
| hua | hua | fua | fua | fruit |
| hulu | huru | fulu | fulu | hair |
| heʔe | $h^w$eke | feʔe | feke | octopus |
| | $h^w$ai | fai | fai | rayfish |
| hiku | $h^w$itu | fitu | fitu | seven |
| haku | $h^w$atu | fatu | fatu | stone |
| niho | niho | nifo | nifo | tooth |
| honu | honu | | fonu | turtle |
| mahana | mahana | mafana | mafana | warm |

# Chapter 9

# Language Acquisition

## Contents


9.1 The Earliest Stages . . . . . . . . . . . . . . . . . . 194

9.2 Later Stages . . . . . . . . . . . . . . . . . . . . . . 197

9.3 Theories of Acquisition . . . . . . . . . . . . . . . . 199

9.4 The Critical Period Hypothesis . . . . . . . . . . . 202

9.5 Summing Up . . . . . . . . . . . . . . . . . . . . . . 203


The process of language acquisition sits at the very center of the major concerns of modern linguistic theory. It is fundamental both to the problem of characterizing the nature of grammar (some part of which is assumed to be an innate endowment of humans, another part learned) and to an understanding of language change (and therefore, of language variation).

While linguistic research has made considerable progress in the study of certain aspects of acquisition, the picture is far from complete. This is partly because children, especially young children, are rather difficult to work with, and partly because experimentation on human subjects is not always possible in general. (Oddly, parents will not generally allow their children to undergo useful experiments which involve procedures like injecting radioactive fluid into their children's brains so that we can get an image of the activity in there, solely in the interest of scientific curiosity.) In addition, children are even worse than adults at conveying anything useful about what's going with them linguistically. In the early stages they can't talk (and we are unable to extract information from bubbles and various gastric noises). Later on, they talk after a fashion, but it turns out they generally have no idea what you're talking about when you ask them perfectly reasonable questions (like: 'Is that an allophone of this?'). Consequently, with the exception of some nifty non-invasive experiments, some for early speech perception and some for comprehension at later stages, most of our data on acquisition comes from the same sources as our data for adult grammars – production and comprehension. For reasons we will examine shortly, using children's production is even more fraught with danger than using adult production and children's comprehension data doesn't lag very far behind.

Before discussing acquisition data and methodology in more detail, we will outline a few points about knowledge of language which are relevant to the question of which aspects of that knowledge are innate, and which aspects must be learned from the environment. An innate portion is suggested by the following:

- Aspects of linguistic knowledge appear to be present at birth.
- The development of language mirrors the development of other, innate abilities.
- The system is too complex to be acquired by children in the time that we know they acquire it.
- The data that children receive is insufficiently robust to allow them to construct a grammatical system without relying on some innate knowledge.

A learned portion is suggested by the very obvious fact that children acquire whatever language (or languages) they are exposed to in their environment. It is not quite as straightforward as simply memorizing vocabulary items, of course, but it seems to be true that a child will be unable to acquire language if he or she is not in some language environment.

## 9.1 The Earliest Stages

While infants are not known for their stimulating conversational repartee, they do demonstrate some surprisingly sophisticated, and apparently language-specific, knowledge of sounds. To study early (from within a few days or even hours of birth to about 4 months) infant speech perception, a particular type of experimental procedure was developed – the High Amplitude Sucking test (HAS). This test takes advantage of one of the few activities babies of this age are capable of as well as of the intriguing fact that babies appear to suck faster on pacifiers when they're interested in or excited by something than when they are bored. In the HAS experiment, a pacifier is attached to a pressure transducer which records sucking rate. A baseline sucking rate is established for each infant and subsequently the infant is exposed to some stimulus. When first exposed to the new stimulus, the sucking rate increases, but as the same stimulus is played repeatedly, the sucking rate drops off (the infant becomes habituated). When the 'old' stimulus is replaced with a new stimulus, the sucking rate increases (and then drops off again after a certain period of time).

Using this technique, experimenters have tested infant perception of a number of differences between natural language speech sounds. For example, when infants are played the sound [ba] in a HAS experiment and then the stimulus is changed to [pa], infants react with increased sucking rate, indicating that they perceive [pa] to be a new stimulus (thus they distinguish between [ba] and [pa]). This same type of experiment has been performed on infants from a range of different language environments and the stimuli have included a range of natural language speech sounds. The results of these experiments suggest that infants in this age range are already sensitive to the possible sound distinctions in language. Moreover, their discrimination, at this early age, is not at all dependent upon the language environment that they have been in. For example, English environment infants are as capable as Hindi environment infants at distinguishing between sounds that are present in Hindi but not in English and vice versa.

Results such as these suggest that the ability to perceive distinctions in natural language speech sounds is innate rather than learned. On the other hand, one might perfectly well argue that all that these experiments show is that infants have very acute hearing and that they are not discriminating between *speech* sounds, merely between *sounds*. In order to try and resolve this question, experimenters have run additional tests based on what we know of adult perception of certain distinctions.

In the HAS experiment described above with [ba]/[pa], the difference that infants are perceiving is one of voicing on the initial consonant. The true description of what 'voicing' consists of for stops is a little more complicated. It turns out that the real cue for the listener is not simply whether or not there is vocal fold vibration but, instead, exactly *when* that vibration starts with respect to when you release your lips (in the case of a bilabial stop). For a listener to perceive a bilabial stop as voiced (i.e. as [ba]), vocal fold vibration must start anywhere between about 10 milliseconds before to 10 milliseconds after the release. In the case of [pa], speakers usually wait until about 20 milliseconds into the [a] before they start voicing (that is, 20 milliseconds after the release). What we have been describing here is *Voice Onset Time* (VOT). VOT is literally when voicing starts with respect to the release of the stop closure (in the case of bilabials, that is when you open your lips). Again, for bilabials, if you start voicing anytime between 10 milliseconds before the stop release and 10 milliseconds after, an adult listener will hear [ba], while if you start voicing sometime after 20 milliseconds after the stop release, adults will hear [pa]. Experimenters used these VOT timings and one additional factor which we will talk about at the same time as the experiment itself.

The experiment proceeds in a similar fashion as the simple [pa]/[ba] experiment but with some fine tuning. First, a series of [ba]'s is played to the infant but the VOT of the [ba]'s is different. The first set contains [ba]'s which start voicing at 10 milliseconds prior to release. When the infant's sucking rate decreases, a second set of [ba]'s which start voicing 10 milliseconds *after* release is played. Critically, the infants show no interest in this 20 millisecond change of voicing – they do not increase sucking rate when the second type is introduced. Similarly if you play a series of [pa]'s, the first set of which has voicing starting 20 milliseconds after the stop release and the second set with voicing starting 40 milliseconds after the stop release, the infant sucking rate continues to fall as if there was no change in stimulus. Note that, in both cases ([ba]'s and [pa]'s), there was a 20 millisecond difference in VOT between the first and second sets.

One more set of stimuli, which differ only by 20 milliseconds, is now added to the picture. The infant is played a [ba] with voicing which starts *simulaneously* with release. Then a second stimulus is introduced in which voicing starts 20 milliseconds after release. In this case, the infant *does* react as if there is a new stimulus and the sucking rate increases. Now in this and the previous two cases, voicing was shifted by exactly 20 milliseconds but the the infant treated two of the 20 millisecond shifts as if nothing new happened unlike the third case. Since the stimili were equidistant from one another, the only explanation for the fact that child treats exactly one pair of them as if they were different is if the child is particularly sensitive to the difference between 0 millisecond delay (after stop release) and a 20 millisecond delay (after stop release) – precisely the difference between a [b] and a [p] in adult language. This phenomenon is called 'categorical perception' and, unlike the first results cited for simple [pa]/[ba], cannot be explained by simply saying that the stimuli were acoustically different. These results seem to indicate something about the presence of knowledge specific to language in infants. To emphasize this, HAS experiments have been done using pure tones, rather than speech sounds. No categorical difference of the type described above exists for the tone-based experiments.[1]

From the production perspective, there is little going on at the earliest ages. Most of the sounds produced by infants up to about 3-5 months fall into the attractively labelled 'vegetative noises' category. After this time, however, infants begin to make sounds which are somewhat more related to speech sounds. At this point, their production is called 'babbling.' The babbling sounds are not linguistically significant in that they are not produced consistently with respect to naming anything, however listeners hear these sounds as speech-like. While babies in this stage produce a very wide variety of speech-like sounds, the majority of the consonant sounds they make are from the set [h d b m t g w n k]. Once again, this set is language environment independent. It seems to have nothing to do with the sounds heard in the particular environment. Given which sounds are most 'popular', it seems likely that babies are producing most often the sounds which require the least sophisticated motor coordination. In general, the babbling stage is felt to be a type of 'practice' for the real thing later on.

## 9.2 Later Stages

From about 12 to 18 months, children start producing some recognizable speech (all things considered).[2] Their earliest utterances consist of single words. As a result, this is called the 'one word stage.' The assumption is that children actually intend something like a sentence but can only manage a little part of it. The latter is due, in our opinion, to general limitations that it is well known children have, such as limited short-term memory, limited

[1] Nothing is ever clear-cut, unfortunately. Chinchillas are also very good at categorical perception experiments.

[2] Note that an awful lot must have gone on to get to this point from the earlier stage. The details of all of this (those that we've managed to figure out, anyway) are too many to include here.

motor coordination, limited attention spans, and so on. If one is limited to one word, a good strategy is to pick the word that will be key in getting your message across – this appears to be what children do.

Interestingly, comprehension far exceeds production at this point in time. In general, production does not catch up to comprehension for a number of years although the gap becomes narrower as time wears on. (One could conceivably make the point that production never actually catches up to comprehension, of course, given adult performance vs. competence.) Children at the one-word stage are able to interpret at least simple sentences relatively well. In a frequently depicted experiment, English environment children at the one-word stage are placed in front of two TV screens. One of the screens will show Big Bird tickling Cookie Monster and the other will show Cookie Monster tickling Big Bird. (Success is predicatd upon familiarity with these characters, obviously.) A loudspeaker will announce something like 'Look! Big Bird is tickling Cookie Monster!' and the children turn toward the appropriate screen and concentrate on it for some time. This shows that the children already have some relatively advanced knowledge of the syntactic structure of English where the subject (the one doing the action) comes first in a simple sentence.

A few months after the one word stage, the next stage in production occurs. This is the 'two word stage' characterized by – you guessed it – utterances of only two words. In this case, the strings are frequently, although not necessarily, composed of a noun and a verb.

- Allgone sock.
- Baby chair.
- Doggie bark.
- No eat.
- Throw ball.

Again, these two word strings are assumed to represent more complete messages (sentences) but due to the limitations mentioned above, only a couple of key words are there. There is no three word stage to follow the two word stage. Instead, some months later, utterances which have varying numbers of words but virtually no function words begin to be produced.

Because this is similar to the style of telegraph messages (where all but the most important words were omitted for reasons of cost), this stage is called the 'telegraphic stage.' This stage might include strings like the following.

- Man ride bus today.
- Me wanna show Mommy.
- Me put it back.
- No do that again.

Beyond this point, no very clear distinctions can be made in terms of stages of production. The number of vocabulary items soars and there is constant progress toward approximating adult speech. By the time children are about 5 years old, their grammars appear little different from adult grammars.

## 9.3 Theories of Acquisition

Over the years, a number of theories have been put forth to account for the fact that children acquire language. One of these theories was the *imitation theory*. It held that children learned language by imitating the people speaking around them. This theory has an obvious disadvantage in that it seems to predict that we should only be able to say things that we've heard others say (if all our knowledge is due to imitation, what else could we do?). As the examples from the two word and telegraphic stages suggest, children are doing something other than imitating people around them. Few adults produce sentences such as 'Man ride bus today' or 'Allgone sock.' Even the more 'subtle' aspects of language development cannot be due to imitation. For example, children's morphological development shows that they deduce the existence of regular, productive morphology first and incorporate the appropriate rules into their grammar. This leads to output such as the following.

| Earlier stage | Later stage |
|---|---|
| dogs | dogs |
| cats | cats |
| mans | men |
| childs | children |
| kicked | kicked |
| goed | went |
| runned | ran |

It is only at a later stage that lexical exceptions are 'recorded' (stored as exceptions in the lexicon) and then produced appropriately. There is no explanation for non-adult forms in the first column under a theory which says that children imitate those around them (i.e. adults).

Another theory, the *reinforcement theory*, popular with behavioral psychologists earlier in this century, held that children made some noise, and if they got positively reinforced by their parents for making this sound, they made it some more. Eventually, through the process of continual adult reinforcement, children came to say words and eventually sentences. An interesting dialogue between a sadistic linguistic parent and her small child reveals some problems with this theory.

Child: Nobody don't like me.

Parent: You mean 'nobody likes me.'

Child: Nobody don't like me.

Parent: Say 'nobody likes me.'

Child: Nobody don't like me.'

Parent: Say 'nobody.'

Child: Nobody.

Parent: Say 'likes.'

Child: Likes.

Parent: Say 'me.'

Child: Me. Nobody don't like me.

Note first that any normal (i.e., non-linguist) parent would respond to this child's pathetic assertion of lack of self-worth with a supportive 'Oh, come on, everybody likes you.' or some such statement. This dialogue thus reveals two difficulties for the 'reinforcement theory.' First, parents generally reinforce children not for the *form* of their assertions but for the *content*. For example, when, as a child, one of the authors said 'Toni is stupid' (Toni is his sister), he did not get a nice pat on the head from his mother because of his proper use of the verb 'to be.' He got, instead, some negative reinforcement in the form of a swat on the behind. On the other hand, when he said 'I wuves you mommy' he got a big hug, in spite of the ill-formedness of his utterance (and the fact that he had his eye on some more candy). Reinforcement will never explain the acquisition of the grammar.

The dialogue is equally bad for the 'imitation theory.' The mother offers several, carefully presented models for the child to imitate and the child stubbornly persists in staying it his way as opposed to imitating. And again, who could the child be imitating when he says 'nobody don't likes me'? This particular child was born into an environment in which there were *no* speakers of any dialect which produced such forms. Instead, what such dialogues seem to reveal is the same aspect of acquisition that the morphology examples revealed. At any particular point, the child has a grammar – it's just a different grammar than an adult grammar. The child's motivation is neither to get some particular reinforcement nor to accurately imitate an adult, but rather to express himself, using the grammar he has constructed.

Regardless of what kind of input they receive, a child learning an English type grammar will produces sentences like 'nobody don't like me.' This can be explained if this type of output reflects a necessary intermediate step between the innate linguistic system, present in all speakers at birth, and the fully-acquired English-type grammar. The *innateness hypothesis*, as this theory is called, has many distinct advantages over its earlier competitors. Theories which have the child starting out with nothing in their heads (so-called *tabula rasa* or 'clean slate' theories) not only can't explain the HAS results, they would also never predict dialogues such as the one cited above. The fact that sentences with this structure are uttered by *every* acquirer of English tells us something important: there is an innate linguistic system. The role of caregivers and others within hearing distance of the child is far less obvious than many people supposed. These speakers do provide the child with all important data but they cannot influence what the child *does* with the data they receive – this appears to be regulated by some unconscious, internal acquisition mechanism of the child.

If we take the facts above together with the obvious complexity of human grammars (as you've seen in this course) and the relatively poor quality of the data that children have to work with (consider all of the performance errors in natural speech), a clear conclusion may be drawn. Children are really good at learning languages because they are preprogammed to do so. They are born with innate knowledge of precisely what kinds of things they might find in a grammar and their learning task is more a matter of weeding out elements they don't need (because it isn't in the particular language they're being exposed to) than a matter of discovering facts of great complexity on the basis of really bad evidence. While the individual lexical items of the language to which they are exposed must be learned, the structural possibilities for all human languages are present in their minds at birth.

## 9.4 The Critical Period Hypothesis

The developmental stages described in this chapter are stages that every human child goes through barring severe and specific neurological abnormalities. While none of the ages are absolute for any one stage, the order of acquisition is the same for all children and the ages at which certain aspects of language develop are approximately the same. In this sense, language acquisition is similar to walking. All children learn to walk in pretty much the same way at pretty much the same time – speeding up the process through outside intervention is not possible, as far as we know.[3] This brings up the interesting question of whether there is a 'critical period' for acquiring language. That is, does language, specifically, one's *first* language, *have* to be acquired within a certain time frame?

There is no humane experiment one can perform in order to help answer this question. The only information that might be relevant comes from a handful of tragic cases in which children have been deprived of linguistic stimulus during the normal developmental years. The most famous and most carefully studied of these cases was of a girl, Genie, who was deliberately deprived of almost all normal human interaction through her first fourteen

[3]It is interesting to speculate on whether a child will walk if they are not surrounded by 'walkers.'

years. A linguist, Susan Curtiss, worked with Genie for some period of time after she was rescued from her home environment. Although Genie made considerable progress, her acquisition was 'spotty' and much more successful in the realm of vocabulary than of structure. Her linguistic development was unlike that of a child from a normal environment and remained far from adult-like. Few conclusions can be drawn about the notion of a critical age for acquisition, though, due to the existence of many other variables such as the impact of psychological trauma and neurological problems. In general, the multitude of factors involved in cases such as these do not allow us to get a clear picture of what might be going on.

## 9.5 Summing Up

Human knowledge of language is a system of great complexity and we are far from understanding all of its intricate mechanisms. On the other hand, how hard can it be if a three year old can do it? (Note that, for a three year old, tying one's shoe is an insurmountable obstacle.) The answer appears to be 'Not hard, if you're human.' but 'Not likely, if you're not.' Humans also seem to treat bipedal locomotion as a trivial activity whereas an attempt to fly by flapping one's arms is doomed to failure. All species do what they are genetically programmed to do. The linguistic system is a property of the human genetic code. Language, in the sense it has in linguistics, is thus a defining property of what it means to be human at the most basic level.