7

Phonology: The Sound Patterns of Language

Speech is human, silence is divine, yet also brutish and dead; therefore we must learn both arts.

THOMAS CARLYLE (1795–1881)

Phonology is the study of telephone etiquette.

A HIGH SCHOOL STUDENT

What do you think is greater: the number of languages in the world, or the number of speech sounds in all those languages? Well, there are thousands of languages, but only hundreds of speech sounds, some of which we examined in the previous chapter. Even more remarkable, only a few dozen features, such as *voicing* or *bilabial* or *stop*, are needed to describe every speech sound that occurs in every human language.

That being the case, why, you may ask, do languages sound so different? One reason is that the sounds form different patterns in different languages. English has nasalized vowels, but only in syllables with nasal consonants. French puts nasal vowels anywhere it pleases, with or without nasal consonants. The speech sound that ends the word *song*—the velar nasal [ŋ]—cannot begin a word in English, but it can in Vietnamese. The common Vietnamese name spelled *Nguyen* begins with this sound, and the reason few of us can pronounce this name correctly is that it doesn't follow the English pattern.

The fact that a sound such as [ŋ] is difficult for an English speaker to pronounce at the beginning of a word, but easy for a Vietnamese speaker, means that there is no general notion of "difficulty of articulation" that can explain all of the sound patterns of particular languages. Rather, the ability to pronounce particular sounds depends on the speaker's unconscious knowledge of the sound patterns of her own language or languages.

The study of how speech sounds form patterns is **phonology**. These patterns may be as simple as the fact that the velar nasal cannot begin a syllable in English, or as complex as why *g* is silent in *sign* but is pronounced in the related word *signature*. To see that this is a pattern and not a one-time exception, just consider the slippery *n* in *autumn* and *autum*nal, or the *b* in *bomb* and *bom*bard.

The word *phonology* refers both to the linguistic knowledge that speakers have about the sound patterns of their language and to the description of that knowledge that linguists try to produce. Thus it is like the way we defined *grammar*: your mental knowledge of your language, or a linguist's description of that knowledge.

Phonology tells you what sounds are in your language and which ones are foreign; it tells you what combinations of sounds could be an actual word, whether it is (*black*) or isn't (*blick*), and what combination of sounds could not be an actual word (*lbick). It also explains why certain phonetic features are important to identifying a word, for example voicing in English as in *pat* versus *bat*, while other features, such as aspiration in English, are not crucial to identifying a word, as we noted in the previous chapter. And it also allows us to adjust our pronunciation of a morpheme, for example the past or plural morpheme, to suit the different phonological contexts that it occurs in, as we will discuss shortly.

In this chapter we'll look at some of the phonological processes that you know, that you acquired as a child, and that yet may initially appear to you to be unreasonably complex. Keep in mind that we are only making explicit what you already know, and its complexity is in a way a wondrous feature of your own mind.

The Pronunciation of Morphemes

> The *t* is silent, as in Harlow.
>
> **MARGOT ASQUITH,** referring to her name being mispronounced by the actress Jean Harlow

Knowledge of phonology determines how we pronounce words and the parts of words we call morphemes. Often, certain morphemes are pronounced differently depending on their context, and we will introduce a way of describing this variation with phonological rules. We begin with some examples from English, and then move on to examples from other languages.

The Pronunciation of Plurals

Nearly all English nouns have a plural form: *cat/cats*, *dog/dogs*, *fox/foxes*. But have you ever paid attention to how plural forms are *pronounced*? Listen to a native speaker of English (or yourself if you are one) pronounce the plurals of the following nouns.

A	B	C	D
cab	cap	bus	child
cad	cat	bush	ox
bag	back	buzz	mouse
love	cuff	garage	criterion
lathe	faith	match	sheep
cam		badge	
can			
call			
bar			
spa			
boy			

The final sound of the plural nouns from Column A is a [z]—a *voiced* alveolar fricative. For column B the plural ending is an [s]—a *voiceless* alveolar fricative. And for Column C it's [əz]. Here is our first example of a morpheme with different pronunciations. Note also that there is a regularity in columns A, B, and C that does not exist in D. The plural forms in D—*children*, *oxen*, *mice*, *criteria*, and *sheep*—are a hodge-podge of special cases that are memorized individually when you acquire English, whether natively or as a second language. This is because there is no way to predict the plural forms of these words.

How do we know how to pronounce this plural morpheme? The spelling, which adds *s* or *es*, is misleading—not a *z* in sight—yet if you know English, you pronounce it as we indicated. When faced with this type of question, it's useful to make a chart that records the phonological environments in which each variant of the morpheme is known to occur. (The more technical term for a variant is **allomorph**.) Writing the words from the first three columns in broad phonetic transcription, we have our first chart for the plural morpheme.

Allomorph	Environment
[z]	After [kæb], [kæd], [bæg], [lʌv], [leð], [kæm], [kæn], [bæŋ], [kɔl], [bar], [spa], [bɔɪ], e.g., [kæbz], [kædz] . . . [bɔɪz]
[s]	After [kæp], [kæt], [bæk], [kʌf], [feθ], e.g., [kæps], [kæts] . . . [feθs]
[əz]	After [bʌs], [bʊʃ], [bʌz], [gəraʒ], [mætʃ], [bædʒ], e.g., [bʌsəz], [bʊʃəz] . . . [bædʒəz]

To discover the pattern behind the way plurals are pronounced, we look for some property of the environment associated with each group of allomorphs. For example, what is it about [kæb] or [lʌv] that determines that the plural morpheme will take the form [z] rather than [s] or [əz]?

To guide our search, we look for **minimal pairs** in our list of words. A minimal pair is two words with different meanings that are identical except for one sound segment that occurs in the same place in each word. For example, *cab* [kæb] and *cad* [kæd] are a minimal pair that differ only in their final segments, whereas *cat* [kæt] and *mat* [mæt] are a minimal pair that differ only in their initial segments. Other minimal pairs in our list include *cap/cab*, *bag/back*, and *bag/badge*.

Minimal pairs whose members take different allomorphs are particularly useful for our search. For example, consider *cab* [kæb] and *cap* [kæp], which respectively take the allomorphs [z] and [s] to form the plural. Clearly, the final segment is responsible, because that is where the two words differ. Similarly for *bag* [bæg] and *badge* [bædʒ]. Their final segments determine the different plural allomorphs [z] and [əz].

Apparently, the distribution of plural allomorphs in English is conditioned by the final segment of the singular form. We can make our chart more concise by considering just the final segment. (We treat diphthongs such as [ɔɪ] as single segments.)

Allomorph	Environment
[z]	After [b], [d], [g], [v], [ð], [m], [n], [ŋ], [l], [r], [a], [ɔɪ]
[s]	After [p], [t], [k], [f], [θ]
[əz]	After [s], [ʃ], [z], [ʒ] , [tʃ], [dʒ]

We now want to understand *why* the English plural follows this pattern. We *always* answer questions of this type by inspecting the *phonetic properties* of the conditioning segments. Such an inspection reveals that the segments that trigger the [əz] plural have in common the property of being *sibilants*. Of the nonsibilants, the *voiceless* segments take the [s] plural, and the *voiced* segments take the [z] plural. Now the rules can be stated in more general terms:

Allomorph	Environment
[z]	After voiced nonsibilant segments
[s]	After voiceless nonsibilant segments
[əz]	After sibilant segments

An even more concise way to express these rules is to assume that the basic or underlying form of the plural morpheme is /z/, with the meaning "plural." This is the "default" pronunciation. The rules tell us when the default does *not* apply:

1. Insert a [ə] before the plural morpheme /z/ when a regular noun ends in a sibilant, giving [əz].
2. Change the plural morpheme /z/ to a voiceless [s] when preceded by a voiceless sound.

These rules will derive the phonetic forms—that is, the pronunciations—of plurals for all regular nouns. Because the basic form of the plural is /z/, if no rule applies, then the plural morpheme will be realized as [z]. The following chart shows how the plurals of *bus*, *butt*, and *bug* are formed. At the top are the basic forms. The two rules apply or not as appropriate as one moves downward. The output of rule 1 becomes the input of rule 2. At the bottom are the phonetic realizations—the way the words are pronounced.

	bus + pl.	*butt* + pl.	*bug* + pl.
Basic representation	/bʌs + z/	/bʌt + z/	/bʌg + z/
Apply rule (1)	ə	NA*	NA
Apply rule (2)	NA	s	NA
Phonetic representation	[bʌsəz]	[bʌts]	[bʌgz]

*NA means "not applicable."

As we have formulated these rules, (1) must apply before (2). If we applied the rules in reverse order, we would derive an incorrect phonetic form for the plural of *bus*, as a diagram similar to the previous one illustrates:

Basic representation	/bʌs + z/
Apply rule (2)	s
Apply rule (1)	ə
Phonetic representation	*[bʌsəs]

The particular phonological rules that determine the phonetic form of the plural morpheme and other morphemes of the language are **morphophonemic rules**. Such rules concern the pronunciation of specific morphemes. Thus the plural morphophonemic rules apply to the plural morpheme specifically, not to all morphemes in English.

Additional Examples of Allomorphs

The formation of the regular past tense of English verbs parallels the formation of regular plurals. Like plurals, some irregular past tenses conform to no particular rule and must be learned individually, such as *go/went*, *sing/sang*, and *hit/hit*. And also like plurals, there are three *phonetic* past-tense morphemes for regular verbs: [d], [t], and [əd]. Here are several examples in broad phonetic transcription. Study sets A, B, and C and try to see the regularity before reading further.

Set A: *gloat* [glot], *gloated* [glotəd]; *raid* [red], *raided* [redəd]

Set B: *grab* [græb], *grabbed* [græbd]; *hug* [hʌg], *hugged* [hʌgd]; *faze* [fez], *fazed* [fezd]; *roam* [rom], *roamed* [romd].

Set C: *reap* [rip], *reaped* [ript]; *poke* [pok], *poked* [pokt]; *kiss* [kɪs], *kissed* [kɪst]; patch [pætʃ], *patched* [pætʃt]

Set A suggests that if the verb ends in a [t] or a [d] (i.e., non-nasal alveolar stops), [əd] is added to form the past tense, similar to the insertion of [əz] to form the plural of nouns that end in sibilants. Set B suggests that if the verb ends in a voiced segment other than [d], you add a voiced [d]. Set C shows us that if the verb ends in voiceless segment other than [t], you add a voiceless [t].

Just as /z/ was the basic form of the plural morpheme, /d/ is the basic form of the past-tense morpheme, and the rules for past-tense formation of regular verbs are much like the rules for the plural formation of regular nouns. These are also *morphophonemic* rules as they apply specifically to the past-tense morpheme /d/. As with the plural rules, the output of Rule 1, if any, provides the input to Rule 2, and the rules must be applied in order.

1. Insert a [ə] before the past-tense morpheme when a regular verb ends in a non-nasal alveolar stop, giving [əd].
2. Change the past-tense morpheme to a voiceless [t] when a voiceless sound precedes it.

Two further allomorphs in English are the possessive morpheme and the third-person singular morpheme, spelled *s* or *es*. These morphemes take on the same phonetic form as the plural morpheme *according to the same rules*! Add [s] to *ship* to get *ship's*; add [z] to *woman* to get *woman's*; and add [əz] to *judge* to get *judge's*. Similarly for the verbs *eat*, *need*, and *rush*, whose third-person singular forms are *eats* with a final [s], *needs* with a final [z], and *rushes* with a final [əz].

That the rules of phonology are based on properties of segments rather than on individual words is one of the factors that makes it possible for young children to learn their native language in a relatively short period. The young child doesn't need to learn each plural, each past tense, each possessive form, and each verb ending, on a noun-by-noun or verb-by-verb basis. Once the rule is learned, thousands of word forms are automatically known. And as we will see when we discuss language development in chapter 8, children give clear evidence of learning morphophonemic rules such as the plural rules by applying the rule too broadly and producing forms such as *mouses*, *mans*, and so on, which are ungrammatical in the adult language.

English is not the only language that has morphemes that are pronounced differently in different phonological environments. Most languages have morpheme variation that can be described by rules similar to the ones we have written for English. For example, the negative morpheme in the West African language Akan has three nasal allomorphs: [m] before *p*, [n] before *t*, and [ŋ] before *k*, as the following examples show ([mɪ] means "I"):

mɪ pɛ	"I like"	mɪ mpɛ	"I don't like"
mɪ tɪ	"I speak"	mɪ ntɪ	"I don't speak"
mɪ kɔ	"I go"	mɪ ŋkɔ	"I don't go"

The rule that describes the distribution of allomorphs is:

> Change the place of articulation of the nasal negative morpheme to agree with the place of articulation of a following consonant.

The rule that changes the pronunciation of nasal consonants as just illustrated is called the **homorganic nasal rule**—*homorganic* means "same place"—because the place of articulation of the nasal is the same as for the following consonant. The homorganic nasal rule is a common rule in the world's languages.

Phonemes: The Phonological Units of Language

> In the physical world the naive speaker and hearer actualize and are sensitive to sounds, but what they feel themselves to be pronouncing and hearing are "phonemes."
>
> **EDWARD SAPIR,** "The Psychological Reality of Phonemes," 1933

The phonological rules discussed in the preceding section apply only to particular morphemes. However, other phonological rules apply to sounds as they occur in any morpheme in the language. These rules express our knowledge about the sound patterns of the entire language.

This section introduces the notions of **phoneme** and **allophone**. Phonemes are what we have been calling the basic form of a sound and are sensed in your mind rather than spoken or heard. Each phoneme has associated with it one or more sounds, called allophones, which represent the actual sound corresponding to the phoneme in various environments. For example, the phoneme /p/ is pronounced with the aspiration allophone [p^h] in *pit* but without aspiration [p] in *spit*. Phonological rules operate on phonemes to make explicit which allophones are pronounced in which environments.

Vowel Nasalization in English as an Illustration of Allophones

English contains a general phonological rule that determines the contexts in which vowels are nasalized. In chapter 6 we noted that both oral and nasal vowels occur *phonetically* in English. The following examples show this:

bean	[bĩn]	bead	[bid]
roam	[rõm]	robe	[rob]

Taking oral vowels as basic—that is, as the phonemes—we have a phonological rule that states:

Vowels are nasalized before a nasal consonant within the same syllable.

This rule expresses your knowledge of English pronunciation: nasalized vowels occur only before nasal consonants and never elsewhere. The effect of this rule is exemplified in Table 7.1.

As the examples in Table 7.1 illustrate, oral vowels in English occur in final position and before non-nasal consonants; nasalized vowels occur only before nasal consonants. The nonwords (starred) show us that nasalized vowels do not occur finally or before non-nasal consonants, nor do oral vowels occur before nasal consonants.

TABLE 7.1 | Nasal and Oral Vowels: Words and Nonwords

Words						Nonwords		
be	[bi]	bead	[bid]	bean	[bĩn]	*[bĩ]	*[bĩd]	*[bin]
lay	[le]	lace	[les]	lame	[lẽm]	*[lẽ]	*[lẽs]	*[lem]
baa	[bæ]	bad	[bæd]	bang	[bæ̃ŋ]	*[bæ̃]	*[bæ̃d]	*[bæŋ]

You may be unaware of this variation in your vowel production, but this is natural. Whether you speak or hear the vowel in *bean* with or without nasalization does not matter. Without nasalization, it might sound a bit strange, as if you had a foreign accent, but *bean* pronounced [bĩn] and *bean* pronounced [bin] would convey the same word. Likewise, if you pronounced *bead* as [bĩd], with a nasalized vowel, someone might suspect you had a cold, or that you spoke nasally, but the word would remain *bead*. Because nasalization is an inessential difference insofar as what the word actually is, we tend to be unaware of it.

Contrast this situation with a change in vowel height. If you intend to say *bead* but say *bad* instead, that makes a difference. The [i] in *bead* and the [æ] in *bad* are sounds from *different* phonemes. Substitute one for another and you get a different word (or no word). The [i] in *bead* and the [ĩ] in the nasalized *bead* do not make a difference in meaning. These two sounds, then, belong to the same phoneme, an abstract high front vowel that we denote between slashes as /i/.

Phonemes are not physical sounds. They are abstract mental representations of the phonological units of a language, the units used to represent words in our mental lexicon. The phonological rules of the language apply to phonemes to determine the pronunciation of words.

The process of substituting one sound for another in a word to see if it makes a difference is a good way to identify the phonemes of a language. Here are twelve words differing only in their vowel:

beat	[bit]	[i]	boot	[but]	[u]
bit	[bɪt]	[ɪ]	but	[bʌt]	[ʌ]
bait	[bet]	[e]	boat	[bot]	[o]
bet	[bɛt]	[ɛ]	bought	[bɔt]	[ɔ]
bat	[bæt]	[æ]	bout	[baʊt]	[aʊ]
bite	[baɪt]	[aɪ]	bot	[bat]	[a]

Any two of these words form a *minimal pair*: two *different* words that differ in one sound. The two sounds that cause the word difference belong to different phonemes. The pair [bid] and [bĩd] are not different words; they are variants of the same word. Therefore, [i] and [ĩ] do *not* belong to different phonemes. They are two actualizations of the same phoneme.

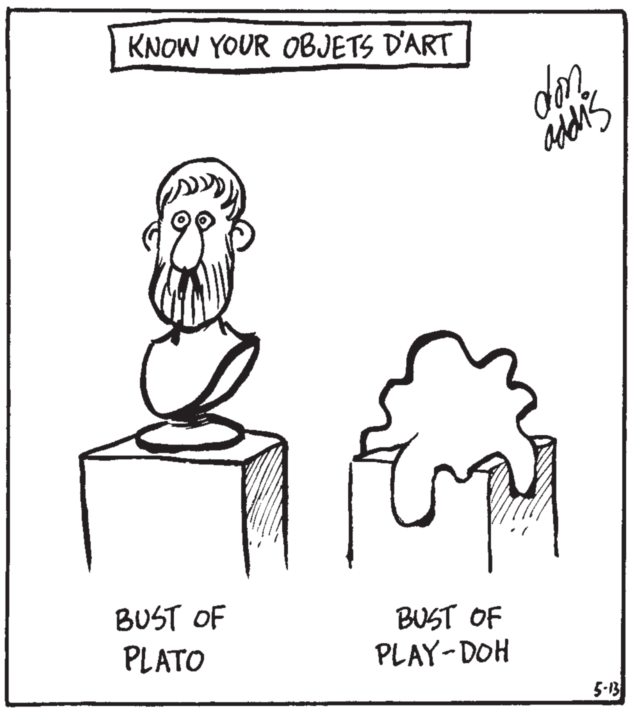
From the minimal set of [b–t] words we can infer that English has at least twelve vowel phonemes. (We consider diphthongs to function as single vowel sounds.) To that total we can add a phoneme corresponding to [ʊ] resulting from minimal pairs such as *book* [bʊk] and *beak* [bik]; and we can add one for [ɔɪ] resulting from minimal pairs such as *boy* [bɔɪ] and *buy* [baɪ].

Our minimal pair analysis has revealed eleven monophthongal and three diphthongal vowel phonemes, namely, /i ɪ e ɛ æ u ʊ o ɔ a ʌ/ and /aɪ/, /aʊ/, /ɔɪ/. (This set may differ slightly in other variants of English.) Importantly, each of these vowel phonemes has (at least) two allophones (i.e., two ways of being pronounced: orally as [i], [ɪ], [e], etc., and nasally as [ĩ], [ɪ̃], [ẽ], etc.), as determined by the phonological rule of nasalization.

A particular realization (pronunciation) of a phoneme is called a **phone**. The collection of phones that are the realizations of the same phoneme are called the *allophones* of that phoneme. In English, each vowel phoneme has both an oral and a nasalized allophone. The choice of the allophone is not random or haphazard; it is *rule-governed*.

To distinguish between a phoneme and its allophones, we use slashes / / to enclose phonemes and continue to use square brackets [ ] for allophones or phones. For example, [i] and [ĩ] are allophones of the phoneme /i/; [ɪ] and [ɪ̃] are allophones of the phoneme /ɪ/, and so on. Thus we will represent *bead* and *bean* phonemically as /bid/ and /bin/. We refer to these as *phonemic* transcriptions of the two words. The rule for the distribution of oral and nasal vowels in English shows that phonetically these words will be pronounced as [bid] and [bĩn]. The pronunciations are indicated by phonetic transcriptions, and written between square brackets.

Allophones of /t/



Consonants, too, have allophones whose distribution is rule-governed. For /t/ the following examples illustrate the point.

tick [tʰɪk] stick [stɪk] hits [hɪts] bitter [bɪɾər]

In *tick* we normally find an aspirated [tʰ], whereas in *stick* and *hits* we find an unaspirated [t], and in *bitter* we find the flap [ɾ]. As with vowel nasalization, swapping these sounds around will not change word meaning. If we pronounce *bitter* with a [tʰ], it will not change the word; it will simply sound unnatural (to most Americans).

We account for this knowledge of how *t* is pronounced by positing a phoneme /t/ with three allophones [tʰ], [t], and [ɾ]. We also posit phonological rules, which roughly state that the aspirated [tʰ] occurs before a stressed vowel, the unaspirated [t] occurs directly before or after /s/, and the flap [ɾ] occurs between a stressed vowel and an unstressed vowel.

Whether we pronounce *tick* as [tʰɪk], [tɪk], or [ɾɪk], we are speaking the same word, however strangely pronounced. The allophones of a phoneme do not *contrast*. If we change the voicing and say *Dick*, or the manner of articulation and say *sick*, or the nasalization and say *nick*, we get different words. Those sounds *do* contrast. *Tick*, *Dick*, *sick*, and *nick* thus form a minimal set that shows us that there are phonemes /t/, /d/, /s/, and /n/ in English. We may proceed in this manner to discover other phonemes by considering *pick*, *kick*, *Mick* (as in Jagger), *Vic*, *thick*, *chick*, *lick*, and *Rick* to infer the phonemes /p/, /k/, /m/, /v/, /θ/, /tʃ/, /l/, and /r/. By finding other minimal pairs and sets, we would discover yet more consonant phonemes such as /ð/, which, together with /θ/, contrasts the words *thy* and *thigh*, or *either* and *ether*.

Each of these phonemes has its own set of allophones, even if that set consists of a single phone, which would mean there is only one pronunciation in all environments. Most phonemes have more than one allophone, and the phonological rules dictate when the different allophones occur. It should be clear at this point that pronunciation is not a random process. It is systematic and rule-governed, and while the systems and the rules may appear complex, they are no more than a compendium of the knowledge that every speaker has.

Complementary Distribution

Minimal pairs illustrate that some speech sounds in a language are contrastive and can be used to make different words such as *big* and *dig*. These contrastive sounds group themselves into the phonemes of that language. Some sounds are non-contrastive and cannot be used to make different words. The sounds [t] and [ɾ] were cited as examples that do not contrast in English, so [raɪtər] and [raɪɾər] are not a minimal pair, but rather alternate ways in which *writer* may be pronounced.

Oral and nasal vowels in English are also non-contrastive sounds. What's more, the oral and nasal allophones of each vowel phoneme never occur in the same phonological context, as Table 7.2 illustrates.

Where oral vowels occur, nasal vowels do not occur, and vice versa. In this sense the phones are said to complement each other or to be in **complementary distribution**. By and large, the allophones of a phoneme are in complementary

TABLE 7.2 | Distribution of Oral and Nasal Vowels in English Syllables

	In Final Position	Before Nasal Consonants	Before Oral Consonants
Oral vowels	Yes	No	Yes
Nasal vowels	No	Yes	No

distribution—never occurring in identical environments. Complementary distribution is a fundamental concept of phonology, and interestingly enough, it shows up in everyday life. Here are a couple of examples that draw on the common experience of reading and writing English.

The first example focuses on *printed* letters such as those that appear on the pages of this book. Each printed letter of English has two main variants: lowercase and uppercase (or capital). If we restrict our attention to words that are not proper names or acronyms (such as Ron or UNICEF), we can formulate a simple rule that does a fair job of determining how letters will be printed:

> A letter is printed in uppercase if it is the first letter of a sentence; otherwise, it is printed in lowercase.

Even ignoring names and acronyms, this rule is only approximately right, but let's go with it anyway. It helps to explain why written sentences such as the following appear so strange:

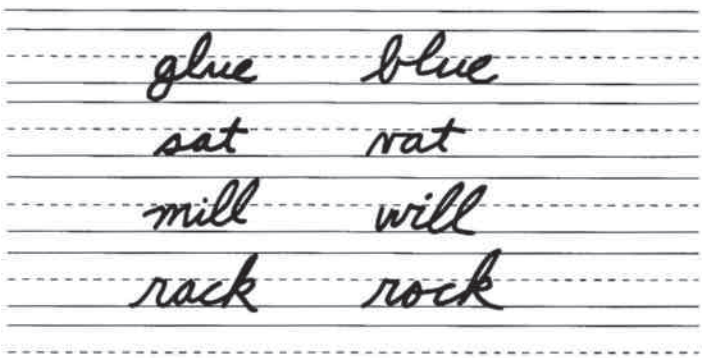
> phonology is the study of the sound patterns of human languageS.
> pHONOLOGY iS tHE sTUDY oF tHE sOUND pATTERNS oF hUMAN lANGUAGES.

These "sentences" violate the rule in funny ways, despite that they are comprehensible, just as the pronunciation of *bead* with a nasal [ĩ] as [bĩd] would sound funny but be understood.

To the extent that the rule is correct, the lowercase and uppercase variants of an English letter *are in complementary distribution*. The uppercase variant occurs in one particular context (namely, at the beginning of the sentence), and the lowercase variant occurs in every other context (or elsewhere). Therefore, just as every English vowel phoneme has an oral and a nasalized allophone that occurs in different spoken contexts, every letter of the English alphabet has two variants, or allographs, that occur in different written contexts. In both cases, the two variants of a single mental representation (phoneme or letter) are in *complementary distribution* because they never appear in the same environment. And, substituting one for the other—a nasal vowel in place of an oral one, or an uppercase letter in place of a lowercase one—may sound or look unusual, but it will not change the meaning of what is spoken or written.

Superman and Clark Kent, or Dr. Jekyll and Mr. Hyde—for those of you familiar with these fictional characters—are in complementary distribution *with respect to time*. At a given moment in time, the individual is either one or another of his alter egos.

Our next example turns to *cursive* handwriting, which you are likely to have learned in elementary school. Writing in cursive is in one sense more similar to the act of speaking than printing is, because in cursive writing each letter of a word (usually) connects to the following letter—just as adjacent sounds connect during speech. The following figure illustrates that the connections between the letters of a word in cursive writing create different variants of a letter in different environments:



Compare how the letter *l* appears after a *g* (as in *glue*) and after a *b* (as in *blue*). In the first case, the *l* begins near the bottom of the line, but in the second case, the *l* begins near the middle of the line (which is indicated by the dashes). In other words, the same letter *l* has two variants. It doesn't matter where the *l* begins, it's still an *l*. Likewise, it doesn't matter whether a vowel in English is nasalized or not, it's still that vowel. Which variant occurs in a particular word is determined by the immediately preceding letter. The variant that begins near the bottom of the line appears after letters like *g* that end near the bottom of the line. The variant that begins near the middle of the line appears after letters like *b* that end near the middle of the line. The two variants of *l* are therefore in complementary distribution.

This pattern of complementary distribution is not specific to *l* but occurs for other cursive letters in English. By examining the pairs *sat* and *vat*, *mill* and *will*, and *rack* and *rock*, you can see the complementary distribution of the variants of *a*, *i*, and *c*, respectively. In each case, the immediately preceding letter determines which variant occurs, with the consequence that the variants of a given letter are in complementary distribution.

We turn now to a general discussion of phonemes and allophones. When sounds are in complementary distribution, they do not contrast with each other. The replacement of one sound for the other will not change the meaning of the word, although it might not sound like typical English pronunciation. Given these facts about the patterning of sounds in a language, a phoneme can be defined as a set of phonetically similar sounds that are in complementary distribution. A set may consist of only one member. Some phonemes are represented by only one sound; they have one allophone. When there is more than one allophone in the set, the phones must be *phonetically similar*; that is, share most phonetic features. In English, the velar nasal [ŋ] and the glottal fricative [h] are in complementary distribution; [ŋ] does not occur word initially and [h] does not occur word finally. But they share very few phonetic features; [ŋ] is a voiced velar nasal stop; [h] is a voiceless glottal fricative. Therefore, they are not allophones of the same phoneme; [ŋ] and [h] are allophones of different phonemes.

Speakers of a language generally perceive the different allophones of a single phoneme as the same sound or phone. For example, most speakers of English are unaware that the vowels in *bead* and *bean* are different phones because mentally, speakers produce and hear phonemes, not phones.

Distinctive Features of Phonemes

We are generally not aware of the phonetic properties or features that distinguish the phonemes of our language. *Phonetics* provides the means to describe the phones (sounds) of language, showing how they are produced and how they vary. *Phonology* tells us how various sounds form patterns to create phonemes and their allophones.

For two phones to contrast meaning, there must be some phonetic difference between them. The minimal pairs *seal* [sil] and *zeal* [zil] show that [s] and [z] represent two contrasting phonemes in English. They cannot be allophones of one phoneme because one cannot replace the [s] with the [z] without changing the meaning of the word. Furthermore, they are not in complementary distribution; both occur word initially before the vowel [i]. They are therefore allophones of the two different phonemes /s/ and /z/. From the discussion of phonetics in chapter 6, we know that [s] and [z] differ in voicing: [s] is voiceless and [z] is voiced. The phonetic feature of voicing therefore distinguishes the two words. Voicing also distinguishes *feel* and *veal* [f]/[v] and *cap* and *cab* [p]/[b]. When a feature distinguishes one phoneme from another, hence one word from another, it is a **distinctive feature** or, equivalently, a **phonemic feature**.

Feature Values

One can think of voicing and voicelessness as the presence or absence of a single feature, *voiced*. This single feature may have two values: plus (+), which signifies its presence, and minus (–), which signifies its absence. For example, [b] is [+voiced] and [p] is [–voiced].

The presence or absence of nasality can similarly be designated as [+nasal] or [–nasal], with [m] being [+nasal] and [b] and [p] being [–nasal]. A [–nasal] sound is an oral sound.

We consider the phonetic and phonemic symbols to be *cover symbols* for sets of distinctive features. They are a shorthand method of specifying the phonetic properties of the segment. Phones and phonemes are not indissoluble units; they are composed of phonetic features, similar to the way that molecules are composed of atoms. A more explicit description of the phonemes /p/, /b/, and /m/ may thus be given in a feature matrix of the following sort.

	p	**b**	**m**
Stop	+	+	+
Labial	+	+	+
Voiced	–	+	+
Nasal	–	–	+

Aspiration is not listed as a phonemic feature in the specification of these units, because it is not necessary to include both [p] and [p^h] as phonemes. In a phonetic transcription, however, the aspiration feature would be specified where it occurs.

A phonetic feature is distinctive when the + value of that feature in certain words contrasts with the – value of that feature in other words. At least one feature value difference must distinguish each phoneme from all the other phonemes in a language.

Because the phonemes /b/, /d/, and /g/ contrast by virtue of their place of articulation features—*labial*, *alveolar*, and *velar*—these place features are also distinctive in English. Because uvular sounds do not occur in English, the place feature *uvular* is not distinctive. The distinctive features of the voiced stops in English are shown in the following:

	b	**m**	**d**	**n**	**g**	**ŋ**
Stop	+	+	+	+	+	+
Voiced	+	+	+	+	+	+
Labial	+	+	–	–	–	–
Alveolar	–	–	+	+	–	–
Velar	–	–	–	–	+	+
Nasal	–	+	–	+	–	+

Each phoneme in this chart differs from all the other phonemes by at least one distinctive feature.

Vowels, too, have distinctive features. For example, the feature [±back] distinguishes the vowel in *rock* [rak] ([+back]) from the vowel in *rack* [ræk] ([–back]), among others, and is therefore distinctive. Similarly, [±tense] distinguishes [i] from [ɪ] (*beat* versus *bit*), among others, and is also a distinctive feature of the vowel system.

Nondistinctive Features

We have seen that nasality is a distinctive feature of English consonants, but it is a **nondistinctive feature** for English vowels. Given the arbitrary relationship between form and meaning, there is no way to predict that the word *meat* begins with a nasal bilabial stop [m] and that the word *beat* begins with an oral bilabial stop [b]. You learn this when you learn the words. On the other hand, the nasality feature value of the vowels in *bean*, *mean*, *comb*, and *sing* is predictable because they occur before nasal consonants. When a feature value is predictable by rule for a certain class of sounds, the feature is a **nondistinctive** or **redundant** or **predictable feature** *for that class*. (The three terms are equivalent.) Thus nasality is a redundant feature in English vowels, but a **nonredundant** (distinctive or phonemic) feature for English consonants.

This is not the case in all languages. In French, nasality is a distinctive feature for both vowels and consonants: *gars* (pronounced [ga]) "lad" contrasts with *gant* [gã], which means "glove"; and *bal* [bal] "dance" contrasts with *mal* [mal] "bad." Thus, French has both oral and nasal consonant phonemes and vowel phonemes; English has oral and nasal consonant phonemes, but only oral vowel phonemes.

Like French, the African language Akan (spoken in Ghana) has nasal vowel phonemes. Nasalization is a distinctive feature for vowels in Akan, as the following examples illustrate:

[ka]	"bite"	[kã]	"speak"
[fi]	"come from"	[fĩ]	"dirty"
[tu]	"pull"	[tũ]	"den"
[nsa]	"hand"	[nsã]	"liquor"
[tʃi]	"hate"	[tʃĩ]	"squeeze"
[pam]	"sew"	[pãm]	"confederate"

Nasalization is not predictable in Akan as it is in English. There is no nasalization rule in Akan, as shown by the minimal pair [pam] and [pãm]. If you substitute an oral vowel for a nasal vowel, or vice versa, you will change the word.

Two languages may have the same phonetic segments (phones) but have two different phonemic systems. Phonetically, both oral and nasalized vowels exist in English and Akan. However, English does not have nasalized vowel phonemes, but Akan does. The same phonetic segments function differently in the two languages. Nasalization of vowels in English is redundant and nondistinctive; nasalization of vowels in Akan is nonredundant and distinctive.

Another nondistinctive feature in English is aspiration. In chapter 6 we pointed out that in English both aspirated and unaspirated voiceless stops occur. The voiceless aspirated stops [pʰ], [tʰ], and [kʰ] and the voiceless unaspirated stops [p], [t], and [k] are in complementary distribution in English, as shown in the following:

Syllable Initial before a Stressed Vowel			**After a Syllable Initial /s/**			**Nonword***		
[pʰ]	[tʰ]	[kʰ]	[p]	[t]	[k]			
pill	*till*	*kill*	*spill*	*still*	*skill*	[pɪl]*	[tɪl]*	[kɪl]*
[pʰɪl]	[tʰɪl]	[kʰɪl]	[spɪl]	[stɪl]	[skɪl]	[spʰɪl]*	[stʰɪl]*	[skʰɪl]*
par	*tar*	*car*	*spar*	*star*	*scar*	[par]*	[tar]*	[kar]*
[pʰar]	[tʰar]	[kʰar]	[spar]	[star]	[skar]	[spʰar]*	[stʰar]*	[skʰar]*

Where the unaspirated stops occur, the aspirated ones do not, and vice versa. If you wanted to, you could say *spit* with an aspirated [pʰ], as [spʰɪt], and it would be understood as *spit*, but listeners would probably think you were spitting out your words. Given this distribution, we see that aspiration is a redundant, nondistinctive feature in English; aspiration is predictable, occurring as a feature of voiceless stops when they occur initially in a stressed syllable.

This is the reason speakers of English usually perceive the [pʰ] in *pill* and the [p] in *spill* to be the same sound, just as they consider the [i] and [ĩ] that represent the phoneme /i/ in *bead* and *bean* to be the same. They do so because the difference between them is *predictable*, *redundant*, *nondistinctive*, and *nonphonemic* (all equivalent terms). This example illustrates why we refer to the phoneme as an abstract unit or as a mental unit. We do not utter phonemes; we produce phones, the allophones of the phonemes of the language. In English /p/ is a phoneme that is realized phonetically (pronounced) as both [p] and [pʰ], depending on context. The phones or sounds [p] and [pʰ] are allophones of the phoneme /p/.

Phonemic Patterns May Vary across Languages

> The tongue of man is a twisty thing, there are plenty of words there of every kind, the range of words is wide, and their variance.
>
> **HOMER,** *The Iliad,* c. 900 B.C.E.

We have seen that the same phones may occur in two languages but pattern differently because the phonologies are different. English, French, and Akan have oral and nasal vowel phones; in English, oral and nasal vowels are allophones of one phoneme, whereas in French and Akan they represent distinct phonemes.

Aspiration of voiceless stops further illustrates the asymmetry of the phonological systems of different languages. Both aspirated and unaspirated voiceless stops occur in English and Thai, but they function differently in the two languages. Aspiration in English is not a distinctive feature because its presence or absence is predictable. In Thai it is not predictable, as the following examples show:

Voiceless Unaspirated		**Voiceless Aspirated**	
[paa]	*forest*	[p^haa]	*to split*
[tam]	*to pound*	[t^ham]	*to do*
[kat]	*to bite*	[k^hat]	*to interrupt*

The voiceless unaspirated and the voiceless aspirated stops in Thai occur in minimal pairs; they contrast and are therefore phonemes. In both English and Thai, the phones [p], [t], [k], [p^h], [t^h], and [k^h] occur. In English they represent the phonemes /p/, /t/, and /k/; in Thai they represent the phonemes /p/, /t/, /k/, /p^h/, /t^h/, and /k^h/. Aspiration is a distinctive feature in Thai; it is a nondistinctive redundant feature in English.

The phonetic facts alone do not reveal what is distinctive or phonemic:

> The *phonetic representation* of utterances shows what speakers know about the pronunciation of sounds.
> The *phonemic representation* of utterances shows what speakers know about the patterning of sounds.

That *pot/pat* and *spot/spat* are phonemically transcribed with an identical /p/ reveals the fact that English speakers consider the [p^h] in *pot* [p^hat] and the [p] in *spot* [spat] to be phonetic manifestations of the same phoneme /p/. This is also reflected in spelling, which is more attuned to phonemes than to individual phones.

In English, vowel length and consonant length are nonphonemic. Prolonging a sound in English will not produce a different word. In other languages, long and short vowels that are identical except for length are phonemic. In such languages, length is a nonpredictable distinctive feature. For example, vowel length is phonemic in Korean, as shown by the following minimal pairs (recall that the colon-like symbol : indicates length):

il	"day"	i:l	"work"
seda	"to count"	se:da	"strong"
kul	"oyster"	ku:l	"tunnel"

In Italian the word for "grandfather" is *nonno* /nonːo/, which contrasts with the word for "ninth," which is *nono* /nono/, so consonant length is phonemic in Italian. In Luganda, an African language, consonant length is also phonemic: /kula/ with a short /k/ means "grow up," whereas /kːula/ with a long /kː/ means "treasure." Thus consonant length is unpredictable in Luganda, just as whether a word begins with a /b/ or a /p/ is unpredictable in English.

ASL Phonology

As discussed in chapter 6, signs can be broken down into smaller units that are in many ways analogous to the phonemes and distinctive features in spoken languages. They can be decomposed into location, movement, and handshape and there are minimal pairs that are distinguished by a change in one or another of these features. Figure 6.6 in chapter 6 provides some examples. The signs meaning "candy," "apple," and "jealous" are articulated at the same location on the face and involve the same movement, but contrast minimally in hand configuration. "Summer," "ugly," and "dry" are a minimal set contrasting only in place of articulation, and "tape," "chair," and "train" contrast only in movement. Thus signs can be decomposed into smaller minimal units that contrast meaning. Some features are non-distinctive. Whether a sign is articulated on the right or left hand does not affect its meaning.

Natural Classes of Speech Sounds

It's as large as life, and twice as natural!

LEWIS CARROLL, *Through the Looking-Glass,* 1871

We show what speakers know about the predictable aspects of speech through phonological rules. In English, these rules determine the environments in which vowels are nasalized or voiceless stops aspirated. These rules apply to *all* the words in the language, and even apply to made-up words such as *sint*, *peeg*, or *sparg*, which would be /sɪnt/, /pig/, and /sparg/ phonemically and [sɪ̃nt], [pʰig], and [sparg] phonetically.

The more linguists examine the phonologies of the world's languages, the more they find that similar phonological rules involve the same classes of sounds such as nasals or voiceless stops. For example, many languages besides English have a rule that nasalizes vowels before nasal consonants:

Nasalize a vowel when it precedes a nasal consonant in the same syllable.

The rule will apply to all vowel phonemes when they occur in a context preceding any segment marked [+nasal] in the same syllable, and will add the feature [+nasal] to the feature matrix of the vowel. Our description of vowel nasalization in English needs only this rule. It need not include a list of the individual vowels to which the rule applies or a list of the sounds that result from its application.

Many languages have rules that refer to [+voiced] and [–voiced] sounds. For example, the aspiration rule in English applies to the class of [–voiced] noncontinuant sounds in word-initial position. As in the vowel nasality rule, we do not need to consider individual segments. The rule automatically applies to initial /p/, /t/, /k/, and /tʃ/.

Phonological rules often apply to **natural classes** of sounds. A natural class is a group of sounds described by a small number of distinctive features such as [–voiced], [–continuant], which describe /p/, /t/, /k/, and /tʃ/. Any individual member of a natural class would require more features in its description than the class itself, so /p/ is not only [–voiced], [–continuant], but also [+labial].

The relationships among phonological rules and natural classes illustrate why segments are to be regarded as bundles of features. If segments were not specified as feature matrices, the similarities among /p/, /t/, /k/ or /m/, /n/, /ŋ/ would be lost. It would be just as likely for a language to have a rule such as

1. Nasalize vowels before *p*, *i*, or *z*.

as to have a rule such as

2. Nasalize vowels before *m*, *n*, or *ŋ*.

Rule 1 has no phonetic explanation, whereas Rule 2 does: the lowering of the velum in anticipation of a following nasal consonant causes the vowel to be nasalized. In Rule 1, the environment is a motley collection of unrelated sounds that cannot be described with a few features. Rule 2 applies to the natural class of nasal consonants, namely sounds that are [+nasal], [+consonantal].

The various classes of sounds discussed in chapter 6 also define natural classes to which the phonological rules of all languages may refer. They also can be specified by + and – feature values. Table 7.3 illustrates how these feature values combine to define some major classes of phonemes. The presence of +/– indicates that the sound may or may not possess a feature depending on its context. For example, word-initial nasals are [–syllabic] but some word-final nasals can be [+syllabic], as in *button* [bʌtn̩].

TABLE 7.3 | Feature Specification of Major Natural Classes of Sounds

Features	Obstruents	Nasals	Liquids	Glides	Vowels
Consonantal	+	+	+	–	–
Sonorant	–	+	+	+	+
Syllabic	–	+/–	+/–	–	+
Nasal	–	+	–	–	+/–

Feature Specifications for American English Consonants and Vowels

Here are feature matrices for vowels and consonants in English. By selecting all segments marked the same for one or more features, you can identify natural classes. For example, the natural class of high vowels /i, ɪ, u, ʊ/ is marked [+high] in the vowel feature chart of Table 7.4; the natural class of voiced stops /b, m, d, n, g, ŋ, dʒ/ are the ones marked [+voice] [–continuant] in the consonant chart of Table 7.5.

TABLE 7.4 | Features of Some American English Vowels

Features	i	ɪ	e	ɛ	æ	u	ʊ	o	ɔ	a	ʌ
High	+	+	−	−	−	+	+	−	−	−	−
Mid	−	−	+	+	−	−	−	+	+	−	+
Low	−	−	−	−	+	−	−	−	−	+	−
Back	−	−	−	−	−	+	+	+	+	+	−
Central	−	−	−	−	−	−	−	−	−	−	+
Round	−	−	−	−	−	+	+	+	+	−	−
Tense	+	−	+	−	−	+	−	+	−	+	−

The Rules of Phonology

But that to come
Shall all be done by the rule.

WILLIAM SHAKESPEARE, *Antony and Cleopatra*, 1623

Throughout this chapter we have emphasized that the relationship between the *phonemic* representation of a word and its *phonetic* representation, or how it is pronounced, is *rule-governed*. Phonological rules are part of a speaker's knowledge of the language.

The phonemic representations are *minimally specified* because some features or feature values are predictable. For example, in English all nasal consonants are voiced, so we don't need to specify voicing in the phonemic feature matrix for nasals. Similarly, we don't need to specify the feature *round* for non–low back vowels. If Table 7.5 was strictly phonemic, then instead of a + in the *voice*-row for *m*, *n*, and *ŋ*, the cells would be left blank, as would the cells in the *round*-row of Table 7.4 for *u*, *ʊ*, *o*, *ɔ*. Such underspecification reflects the redundancy in the phonology, which is also part of a speaker's knowledge of the sound system. The phonemic representation should include only the nonpredictable, distinctive features of the phonemes in a word. The phonetic representation, derived by applying the phonological rules, includes all of the linguistically relevant phonetic aspects of the sounds. It does not include all of the physical properties of the sounds of an utterance, however, because the physical signal may vary in many ways that have little to do with the phonological system. The absolute pitch of the sound, the rate of speech, or its loudness is not linguistically significant. The phonetic transcription is therefore also an abstraction from the physical signal; it includes the nonvariant phonetic aspects of the utterances, those features that remain relatively constant from speaker to speaker and from one time to another.

Although the specific rules of phonology differ from language to language, the kinds of rules, what they do, and the natural classes they refer to are universal.

Assimilation Rules

We have seen that nasalization of vowels in English is nonphonemic because it is predictable by rule. The vowel nasalization rule is an assimilation rule, or a rule

TABLE 7.5 | Features of Some American English Consonants

Features	p	b	m	t	d	n	k	g	ŋ	f	v	θ	ð	s	z	ʃ	ʒ	tʃ	dʒ	l	r	j	w	h
Consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	−	−
Sonorant	−	−	+	−	−	+	−	−	+	−	−	−	−	−	−	−	−	−	−	+	+	+	+	+
Syllabic	−	−	−/+	−	−	−/+	−	−	−/+	−	−	−	−	−	−	−	−	−	−	−/+	−/+	−	−	−
Nasal	−	−	+	−	−	+	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Voiced	−	+	+	−	+	+	−	+	+	−	+	−	+	−	+	−	+	−	+	+	+	+	+	−
Continuant	−	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+	−	−	+	+	+	+	+
Labial	+	+	+	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	−	−	−	−	+	−
Alveolar	−	−	−	+	+	+	−	−	−	−	−	−	−	+	+	−	−	−	−	+	+	−	−	−
Palatal	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+	−	−	+	−	−
Anterior	+	+	+	+	+	+	−	−	−	+	+	+	+	+	+	−	−	−	−	+	+	−	−	−
Velar	−	−	−	−	−	−	+	+	+	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−
Coronal	−	−	−	+	+	+	−	−	−	−	−	+	+	+	+	+	+	+	+	+	+	+	−	−
Sibilant	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+	+	+	−	−	−	−	−

Note: The phonemes /r/ and /l/ are distinguished by the feature [lateral], not shown here. /l/ is the only phoneme that would be [+lateral].

that makes neighboring segments more similar by duplicating a phonetic property. For the most part, assimilation rules stem from articulatory processes. There is a tendency when we speak to increase the ease of articulation. It is easier to lower the velum while a vowel is being pronounced before a nasal stop than to wait for the completion of the vowel and then require the velum to move suddenly.

We now wish to look more closely at the phonological rules we have been discussing. Previously, we stated the vowel nasalization rule:

Vowels are nasalized before a nasal consonant within the same syllable.

This rule specifies the <u>class of sounds</u> affected by the rule:

Vowels

It states what <u>phonetic change</u> will occur by applying the rule:

Change phonemic oral vowels to phonetic nasal vowels.

And it specifies the context or <u>phonological environment</u>.

Before a nasal consonant within the same syllable.

A shorthand notation to write rules, similar to the way scientists and mathematicians use symbols, makes the rule statements more concise. Every physicist knows that $E = mc^2$ means "Energy equals mass times the square of the velocity of light." We can use similar notations to state the nasalization rule as:

V → [+nasal] / __ [+nasal] $

Let's look at the rule piece by piece.

V	→	[+nasal]	/	__	[+nasal]	$
Vowels	become	nasalized	in the environment	before	nasal segments	within a syllable

To the left of the arrow is the <u>class of sounds</u> that is affected. To the right of the arrow is the <u>phonetic change</u> that occurs. <u>The phonological environment</u> follows the slash. The underscore __ is the relative position of the sound to be changed within the environment, in this case *before* a nasal segment. The dollar sign denotes a syllable boundary and guarantees that the environment does not cross over to the next syllable.

This rule tells us that the vowels in such words as *den* /dɛn/ will become nasalized to [dɛ̃n], but *deck* /dɛk/ will not be affected and is pronounced [dɛk] because /k/ is not a nasal consonant. As well, a word such as *den$tal* /dɛn$təl/ will be pronounced [dɛ̃n$təl], where we have showed the syllable boundary explicitly. However, the first vowel in *de$note*, /di$not/, will not be nasalized, because the nasal segment does not precede the syllable boundary, so the "within a syllable" condition is not met.

Any rule written in formal notation can be stated in words. The use of formal notation is a shorthand way of presenting the information. Notation also reveals the *function* of the rule more explicitly than words. It is easy to see in the formal statement of the rule that this is an assimilation rule because the change to [+nasal] occurs before [+nasal] segments. Assimilation rules in languages reflect **coarticulation**—the spreading of phonetic features either in the anticipation or in the perseveration (the "hanging on") of articulatory processes. The auditory effect is that words sound smoother.

The following example illustrates how the English vowel nasalization rule applies. It also shows the assimilatory nature of the rule, that is, the change from no nasal feature to [+nasal]:

	"bob"			"boom"		
Phonemic representation	/b	a	b/	/b	u	m/
Nasality: phonemic feature value	–	0*	–	–	0	+
Apply nasal rule		NA			↓	
Nasality: phonetic feature value	–	–	–	–	+	+
Phonetic representation	[b	a	b]	[b	ũ	m]

*The 0 means not present on the phonemic level.

There are many assimilation rules in English and other languages. Recall that the voiced /z/ of the English regular plural suffix is changed to [s] after a voiceless sound, and that similarly the voiced /d/ of the English regular past-tense suffix is changed to [t] after a voiceless sound. These are instances of voicing assimilation. In these cases the value of the voicing feature goes from [+voice] to [–voice] because of assimilation to the [–voice] feature of the final consonant of the stem, as in the derivation of *cats*:

/kæt + z/ → [kæts]

We saw a different kind of assimilation rule in Akan, where we observed that the nasal negative morpheme was expressed as [m] before /p/, [n] before /t/, and [ŋ] before /k/. (This is the homorganic nasal rule.) In this case the place of articulation—bilabial, alveolar, velar—of the nasal assimilates to the place of articulation of the following consonant. The same process occurs in English, where the negative morpheme prefix spelled *in-* or *im-* agrees in place of articulation with the word to which it is prefixed, so we have *impossible* [ĩmp^{h}asəbəl], *intolerant* [ĩnt^{h}alərə̃nt], and *incongruous* [ĩŋk^{h}ãngruəs]. In effect, the rule makes two consonants that appear next to each other more similar.

ASL and other signed languages also have assimilation rules. One example is handshape assimilation, which takes place in compounds such as the sign for "blood." This ASL sign is a compound of the signs for "red" and "flow." The handshape for "red" alone is formed at the chin by a closed hand with the index finger pointed up. In the compound "blood" this handshape is replaced by that of the following word "flow," which is an open handshape (all fingers extended). In other words, the handshape for "red" has undergone assimilation. The location of the sign (at the chin) remains the same. Examples such as this tell us that while the features of signed languages are different from those of spoken languages, their phonologies are organized according to principles like those of spoken languages.

Dissimilation Rules

It is understandable that so many languages have assimilation rules; they permit greater ease of articulation. It might seem strange, then, to learn that languages also have **dissimilation rules**, in which a segment becomes less similar to another segment. Ironically, such rules have the same explanation: it is sometimes easier to articulate dissimilar sounds. The difficulty of tongue twisters like "the sixth sheik's sixth sheep is sick" is based on the repeated similarity of sounds. If one were to make some sounds less similar, as in "the second sheik's tenth sheep is sick," it would be easier to say. The cartoon makes the same point, with *toy boat* being more difficult to articulate repeatedly than *sail boat*, because the [ɔɪ] of *toy* is more similar to [o] than is the [e] of *sail*.

An example of easing pronunciation through dissimilation is found in some varieties of English, where there is a fricative dissimilation rule. This rule applies to sequences /fθ/ and /sθ/, changing them to [ft] and [st]. Here the fricative /θ/ becomes dissimilar to the preceding fricative by becoming a stop. For example, the words *fifth* and *sixth* come to be pronounced as if they were spelled *fift* and *sikst*.

A classic example of the same kind of dissimilation occurred in Latin, and the results of this process show up in the derivational morpheme /-ar/ in English. In Latin a derivational suffix *-alis* was added to nouns to form adjectives. When the suffix was added to a noun that contained the liquid /l/, the suffix was changed to *-aris*; that is, the liquid /l/ was changed to the dissimilar liquid /r/. These words came into English as adjectives ending in *-al* or in its dissimilated form *-ar*, as shown in the following examples:

-al	**-ar**
anecdot-al	angul-ar
annu-al	annul-ar
ment-al	column-ar
pen-al	perpendicul-ar
spiritu-al	simil-ar
ven-al	vel-ar

All of the *-ar* adjectives contain an /l/, and as *columnar* illustrates, the /l/ need not be the consonant directly preceding the dissimilated segment.

Though dissimilation rules are rarer than assimilation rules, they are nevertheless found throughout the world's languages.

Feature-Changing Rules

The assimilation and dissimilation rules we have seen may all be thought of as *feature-changing rules*. In some cases a feature already present is changed. The /z/ plural morpheme has its voicing value changed from plus to minus when it follows a voiceless sound. Similarly, the /n/ in the phonemic negative prefix morpheme /ɪn/ undergoes a change in its place of articulation feature when preceding bilabials or velars. In the case of the Latin dissimilation rule, the feature [+lateral] is changed to [–lateral], so that /l/ is pronounced [r].

The addition of a feature is the other way in which we have seen features change. The English vowel nasalization rule is a case in point. Phonemically, vowels are not marked for nasality; however, in the environment specified by the rule, the feature [+nasal] is added.

Some feature-changing rules are neither assimilation nor dissimilation rules. The rule in English that aspirates voiceless stops at the beginning of a syllable simply adds a nondistinctive feature. Generally, aspiration occurs only if the following vowel is stressed. The /p/ in *pit* and *repeat* is an aspirated [p^h], but the /p/ in *inspect* or *compass* is an unaspirated [p]. We also note that even with an intervening consonant, the aspiration takes place so that words such as *crib*, *clip*, and *quip* ([kʰrɪb], [kʰlɪp], and [kʰwɪp]) all begin with an aspirated [kʰ]. And finally, the affricate /tʃ/ is subject to the rule, so *chip* is phonetically [tʃʰɪp]. We can now state the rule:

> A voiceless, noncontinuant has [+aspirated] added to its feature matrix at the beginning of a syllable containing a stressed vowel with an optional intervening consonant.

Aspiration is not specified in any phonemic feature matrices of English. The aspiration rule adds this feature for reasons having to do with the timing of the closure release rather than in an attempt to make segments more alike or not alike, as with assimilation and dissimilation rules.

Remember that /p/ and /b/ (and all such symbols) are simply cover symbols that do not reveal the phonemic distinctions. In phonemic and phonetic feature matrices, these differences are made explicit, as shown in the following phonemic matrices:

	p	b	
Consonantal	+	+	
Continuant	–	–	
Labial	+	+	
Voiced	–	+	← distinctive difference

The nondistinctive feature "aspiration" is not included in these phonemic representations because aspiration is predictable.

Segment Insertion and Deletion Rules

Phonological rules may add or delete entire segments. These are different from the feature-changing and feature-adding rules we have seen so far, which affect only parts of segments. The process of inserting a consonant or vowel is called **epenthesis**.

The rules for forming regular plurals, possessive forms, and third-person singular verb agreement in English all require an epenthesis rule. Here is the first part of that rule that we gave earlier for plural formation:

> Insert a [ə] before the plural morpheme /z/ when a regular noun ends in a sibilant, giving [əz].

Letting the symbol ∅ stand for "null," we can write this *morphophonemic* epenthesis rule more formally as "null becomes schwa between two sibilants," or like this:

∅ → ə / [+sibilant] ___ [+sibilant]

Similarly, we recall the first part of the rule for regular past-tense formation in English:

> Insert a [ə] before the past-tense morpheme when a regular verb ends in a non-nasal alveolar stop, giving [əd].

This epenthesis rule may also be expressed in our more formal notation:

∅ → ə / [– nasal, + alveolar, – continuant] ___ [– nasal, + alveolar, – continuant]

There is a plausible explanation for insertion of a [ə]. If we merely added a [z] to *squeeze* to form its plural, we would get [skwizː], which would be hard for English speakers to distinguish from [skwiz]. Similarly, if we added just [d] to *load* to form its past tense, it would be [lodː], which would also be difficult to distinguish from [lod], because in English we do not contrast long and short consonants. These and other examples suggest that the morphological patterns in a language are closely related to other generalizations about the phonology of that language.

Just as vowel length can be used for emphasis without changing the meaning of a word, as in "Stooooop [staːp] hitting me," an epenthetic schwa can have a similar effect, as in "P-uh-lease [p^{h}əliz] let me go."

Segment deletion rules are commonly found in many languages and are far more prevalent than segment insertion rules. One such rule occurs in casual or rapid speech. We often delete the unstressed vowels that are shown in bold type in words like the following:

myst**e**ry gen**e**ral mem**o**ry fun**e**ral vig**o**rous Barb**a**ra

These words in casual speech sound as if they were written:

mystry genral memry funral vigrous Barbra

The silent *g* that torments spellers in such words as *sign* and *design* is actually an indication of a deeper phonological process, in this case, one of segment deletion. Consider the following examples:

A		**B**	
sign	[sãɪn]	signature	[sɪgnətʃər]
design	[dəzãɪn]	designation	[dɛzɪgneʃə̃n]
paradigm	[p^{h}ærədãɪm]	paradigmatic	[p^{h}ærədɪgmærək]

In none of the words in column A is there a phonetic [g], but in each corresponding word in column B a [g] occurs. Our knowledge of English phonology accounts for these phonetic differences. The "[g]—no [g]" alternation is regular, and we apply it to words that we never have heard. Suppose someone says:

"He was a salignant [səlɪgnə̃nt] man."

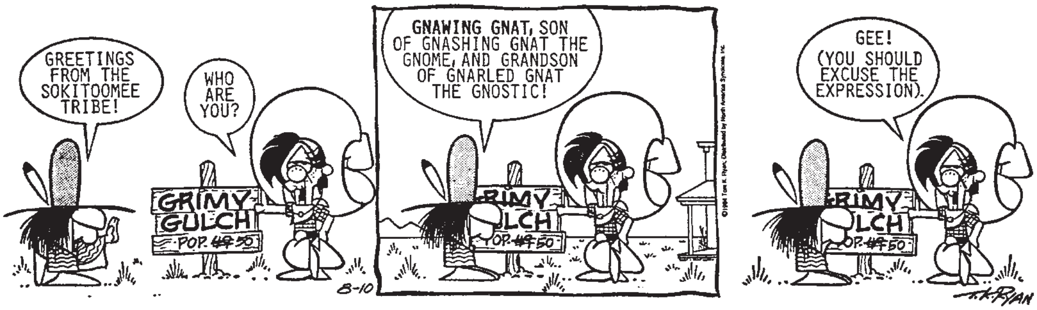
Not knowing what the word means (which you couldn't, since we made it up), you might ask:

"Why, did he salign [səlãɪn] somebody?"

It is highly doubtful that a speaker of English would pronounce the verb form without the *-ant* as [səlɪgn], because the phonological rules of English would delete the /g/ when it occurred in this context. This rule might be stated as:

Delete a /g/ when it occurs before a syllable-final nasal consonant.

The rule is even more general, as evidenced by the pair *gnostic* [nastɪk] and *agnostic* [ægnastɪk], and by the silent *g*'s in the cartoon:



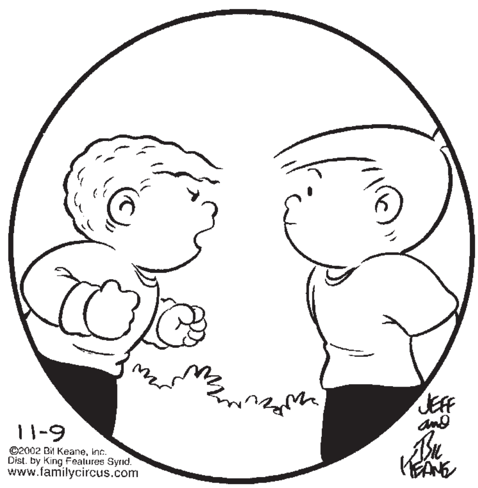
"Tumbleweeds" © Tom K. Ryan. Reprinted with permission of North America Syndicate.

This more general rule may be stated as:

Delete a /g/ word initially before a nasal consonant or before a syllable-final nasal consonant.

Given this rule, the phonemic representation of the stems in *sign/signature*, *design/ designation*, *malign/malignant*, *phlegm/phlegmatic*, *paradigm/paradigmatic*, *gnostic/agnostic*, and so on will include a /g/ that will be deleted by the regular rule if a prefix or suffix is not added. By stating the class of sounds that follow the /g/ (nasal consonants) rather than any specific nasal consonant, the rule deletes the /g/ before both /m/ and /n/.

Movement (Metathesis) Rules



"The only reason I say 'aminal' is I can't say 'animal'!"

"Family Circus" © Bil Keane, Inc. Reprinted with permission of King Features Syndicate.

Phonological rules may also reorder sequences of phonemes, in which case they are called **metathesis** rules. For some speakers of English, the word *ask* is pronounced [æks], but the word *asking* is pronounced [æskĩŋ]. In this case a metathesis rule reorders the /s/ and /k/ in certain contexts. In Old English the verb was *aksian*, with the /k/ preceding the /s/. A historical metathesis rule switched these two consonants, producing *ask* in most dialects of English. Children's speech shows many cases of metathesis (which are corrected as the child approaches the adult grammar): *aminal* [æ̃mə̃nəl] for *animal* and *pusketti* [p^həskɛti] for *spaghetti* are common children's pronunciations. Dog lovers have metathesized the She*tl*and sheepdog into a she*lt*ie, and at least two presidents of the United States have applied a metathesis rule to the word *nuclear*, which many Americans pronounce [n**j**u**k**liər], but is pronounced [nu**k**jələr] by those leading statesmen.

From One to Many and from Many to One

As we've seen, phonological rules that relate phonemic to phonetic representations have several functions, among which are the following:

Function	Example
1. Change feature values	Nasal consonant assimilation rules in Akan and English
2. Add new features	Aspiration in English
3. Delete segments	*g*-deletion before nasals in English
4. Add segments	Schwa insertion in English plural and past tense
5. Reorder segments	Metathesis rule relating [æsk] and [æks]

The relationship between the phonemes and phones of a language is complex and varied. Rarely is a single phoneme realized as one and only one phone. We often find one phoneme realized as several phones, as in the case with English voiceless stops that may be realized as aspirated or unaspirated, among other possibilities. And we find the same phone may be the realization of several different phonemes. Here is a dramatic example of that many-to-one relationship.

Consider the vowels in the following pairs of words:

	A		B	
/i/	comp**e**te	[i]	comp**e**tition	[ə]
/ɪ/	med**i**cinal	[ɪ]	med**i**cine	[ə]
/e/	maint**ai**n	[e]	maint**e**nance	[ə]
/ɛ/	t**e**legraph	[ɛ]	t**e**legraphy	[ə]
/æ/	an**a**lysis	[æ]	an**a**lytic	[ə]
/a/	s**o**lid	[a]	s**o**lidity	[ə]
/o/	ph**o**ne	[o]	ph**o**netic	[ə]
/ʊ/	Talm**u**dic	[ʊ]	Talm**u**d	[ə]

In column A all the boldfaced vowels are stressed vowels with a variety of vowel phones; in column B the boldfaced vowels are without stress or **reduced** and are pronounced as schwa [ə]. In these cases the stress pattern of the word varies because of the different suffixes. The vowel that is stressed in one form becomes reduced in a different form and is therefore pronounced as [ə]. The phonemic representations of all of the root morphemes contain an unreduced vowel such as /i/ or /e/ that is phonetically [ə] when it is reduced. We can conclude, then, that [ə] is an allophone of all English vowel phonemes. The rule to derive the schwa is simple to state:

Change a vowel to a [ə] when the vowel is reduced.

In the phonological description of a language, it is not always straightforward to determine phonemic representations from phonetic transcriptions. How would we deduce the /o/ in *phonetic* from its pronunciation as [fə̃nɛɾɪk] without a complete phonological analysis? However, given the phonemic representation and the phonological rules, we can always derive the correct phonetic representation. In our internal mental grammars this derivation is no problem, because the words occur in their phonemic forms in our mental lexicons and we know the rules of the language.

Similar rules exist in other languages that show that there is no one-to-one relationship between phonemes and phones. For example, in German both voiced and voiceless obstruents occur as phonemes, as is shown by the following minimal pair:

Tier [tiːr] "animal" *dir* [diːr] "to you"

However, when voiced obstruents occur at the end of a word or syllable, they become voiceless. The words meaning "bundle" *Bund* /bʊnd/ and "colorful" *bunt* /bʊnt/ are phonetically identical and pronounced [bʊnt] with a final [t]. Obstruent voicing is neutralized in syllable-final position.

The German devoicing rule changes the specifications of features. In German, the phonemic representation of the final stop in *Bund* is /d/, specified as [+voiced]; it is changed by rule to [–voiced] to derive the phonetic [t] in word-final position. Again, this shows there is no simple relationship between phonemes and their allophones. German presents us with this picture:

$$\begin{tikzcd} \text{German Phonemes} & /d/ \arrow[d] \arrow[dr] & /t/ \arrow[d] \\ \text{German Phones} & [d] & [t] \end{tikzcd}$$

The devoicing rule in German provides a further illustration that we cannot discern the phonemic representation of a word given only the phonetic form; [bʊnt] can be derived from either /bʊnd/ or /bʊnt/. The phonemic representations and the phonological rules together determine the phonetic forms.

The Function of Phonological Rules

The function of the phonological rules in a grammar is to provide the phonetic information necessary for the pronunciation of utterances. We may illustrate this point in the following way:

input *Phonemic (Mental Lexicon) Representation of Words in a Sentence*

↓

Phonological rules (P-rules)

↓

output *Phonetic Representation of Words in a Sentence*

The input to the P-rules is the phonemic representation. The P-rules apply to the phonemic strings and produce as output the phonetic representation.

The application of rules in this way is called a **derivation**. We have given examples of derivations that show how plurals are derived, how phonemically oral vowels become nasalized, and how /t/ and /d/ become flaps in certain environments. A derivation is thus an explicit way of showing both the effects and the function of phonological rules in a grammar.

All the examples of derivations we have so far considered show the application of just one phonological rule, except the plural and past-tense rules, which are actually one rule with two parts. In any event, it is common for more than one rule to apply to a word. For example, the word *tempest* is phonemically /tɛmpɛst/ (as shown by the pronunciation of *tempestuous* [t^{h}ɛ̃mp^{h}ɛstʃuəs]) but phonetically [t^{h}ɛ̃mpəst]. Three rules apply to it: the aspiration rule, the vowel nasalization rule, and the schwa rule. We can derive the phonetic form from the phonemic representation as follows:

Underlying phonemic representation	/	t	ɛ	m	p	ɛ	s	t	/
Aspiration rule		t^h							
Nasalization rule			ɛ̃						
Schwa rule						ə			
Surface phonetic representation	[	t^h	ɛ̃	m	p	ə	s	t	]

Slips of the Tongue: Evidence for Phonological Rules

Slips of the tongue, or **speech errors**, in which we deviate in some way from the intended utterance, show phonological rules in action. We all make speech errors, and they tell us interesting things about language and its use. Consider the following speech errors:

	Intended Utterance	Actual Utterance
1.	gone to seed	god to seen
	[gãn tə sid]	[gad tə sĩn]
2.	stick in the mud	smuck in the tid
	[stɪk ĩn ðə mʌd]	[smʌk ĩn ðə t^{h}ɪd]
3.	speech production	preach seduction
	[spitʃ p^{h}rədʌkʃə̃n]	[p^{h}ritʃ sədʌkʃə̃n]

In the first example, the final consonants of the first and third words were reversed. Notice that the reversal of the consonants also changed the nasality of the vowels. The vowel [ã] in the intended utterance is replaced by [a]. In the actual utterance, the nasalization was lost because it no longer occurred before a nasal consonant. The vowel in the third word, which was the non-nasal [i] in the intended utterance, became [ĩ] in the error, because it was followed by /n/. The nasalization rule applied.

In the other two errors, we see the application of the aspiration rule. In the intended *stick*, the /t/ would have been realized as an unaspirated [t] because it follows the syllable initial /s/. When it was switched with the /m/ in *mud*, it was pronounced as the aspirated [t^h], because it occurred initially. The third example also illustrates the aspiration rule in action. More than being simply amusing, speech errors are linguistically interesting because they provide further evidence for phonological rules and for the decomposition of speech sounds into features.

We will learn more about speech errors in chapter 9 on language processing.

Prosodic Phonology

Syllable Structure

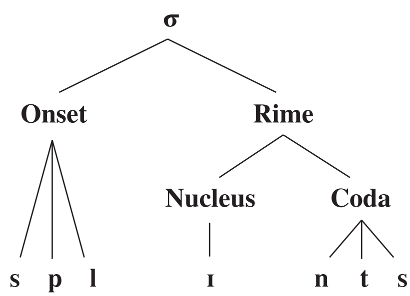
Baby Blues © Baby Blues Partnership. King Features Syndicate

Words are composed of one or more syllables. A **syllable** is a phonological unit composed of one or more phonemes. Every syllable has a **nucleus**, which is usually a vowel (but which may be a syllabic liquid or nasal). The nucleus may be preceded and/or followed by one or more phonemes called the syllable **onset** and **coda**. From a very early age, children learn that certain words rhyme. In rhyming words, the nucleus and the coda of the final syllable of both words are identical, as in the following jingle:

> Jack and **Jill**
> Went up the h**ill**
> To fetch a pail of water.
> Jack fell d**own**
> And broke his cr**own**
> And Jill came tumbling after.

For this reason, the nucleus + coda constitute the subsyllabic unit called a **rime** (note the spelling).

A syllable thus has a hierarchical structure. Using the IPA symbol σ for the phonological syllable, the hierarchical structure of the monosyllabic word *splints* can be shown:



Word Stress

In many languages, including English, one or more of the syllables in every content word (i.e., every word except for function words like *to*, *the*, *a*, *of*) are stressed. A stressed syllable, which can be marked by an acute accent (´), is perceived as more prominent than an unstressed syllable, as shown in the following examples:

pérvert	(noun)	as in	“My neighbor is a pervert.”
pervért	(verb)	as in	“Don’t pervert the idea.”
súbject	(noun)	as in	“Let’s change the subject.”
subjéct	(verb)	as in	“He’ll subject us to criticism.”

These pairs show that stress can be contrastive in English. In these cases it distinguishes between nouns and verbs.

Some words may contain more than one stressed vowel, but exactly one of the stressed vowels is more prominent than the others. The vowel that receives primary stress is marked by an acute accent. The other stressed vowels are indicated by a grave accent (`) over the vowels (these vowels receive secondary stress).

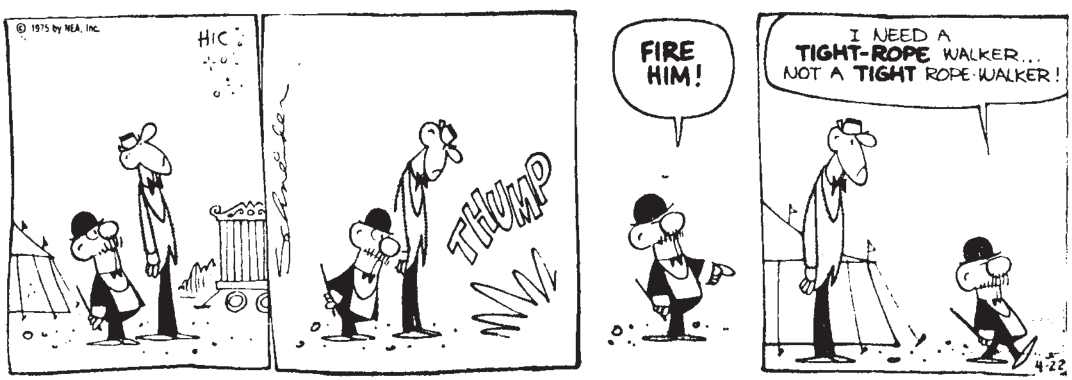
rèsignátion	lìnguístics	sỳstəmátic
fùndaméntal	ìntrodúctory	rèvolútion

Generally, speakers of a language know which syllable receives primary stress, which ones receive secondary stress, and which ones are reduced (are unstressed). It is part of their implicit knowledge of the language. It's usually easy to distinguish between stressed and reduced syllables, because the vowel in reduced syllables is pronounced as a schwa [ə], except at the ends of certain words such as *confetti* or *laboratory*. It may be harder to distinguish between primary and secondary stress. If you are unsure of where the primary stress is in a word (and you are a native or near-native speaker of English), try shouting the word as if talking to a person across a busy street. Often, the difference in stress becomes more apparent.

The stress pattern of a word may differ among English-speaking people. For example, in most varieties of American English the word *láboratòry* [lǽbərət^{h}ɔ̀ri] has two stressed syllables, but in most varieties of British English it receives only one stress [ləbɔ́rətri]. Because English vowels generally reduce to schwa or delete when they are not stressed, the British and American vowels differ in this word. In fact, in the British version the fourth vowel is deleted because it is not stressed.

Stress is a property of the syllable rather than a segment; it is a prosodic or suprasegmental feature. To produce a stressed syllable, one may change the pitch (usually by raising it), make the syllable louder, or make it longer. We often use all three of these phonetic means to stress a syllable.

Sentence and Phrase Stress



"Bimbo's Circus" © Howie Schneider/Dist. by Newspaper Enterprise Association, Inc.

When words are combined into phrases and sentences, one syllable receives greater stress than all others. That is, just as there is only one primary stress in a word spoken in isolation, only one of the vowels in a phrase (or sentence) receives primary stress or accent. All of the other stressed vowels are reduced to secondary stress. In English we place primary stress on the adjectival part of a compound noun (which may be written as one word, two words separated by a hyphen, or two separate words), but we place the stress on the noun when the words are a noun phrase consisting of an adjective followed by a noun. The differences between the following pairs are therefore predictable:

Compound Noun	**Adjective + Noun**
tíghtrope ("a rope for acrobatics")	tight rópe ("a rope drawn taut")
Rédcoat ("a British soldier")	red cóat ("a coat that is red")
hótdog ("a frankfurter")	hot dóg ("an overheated dog")
Whíte House ("the President's house")	white hóuse ("a house painted white")

Say these examples out loud, speaking naturally, and at the same time listen or feel the stress pattern. If English is not your native language, listen to a native speaker say them.

These pairs show that stress may be predictable from the morphology and syntax. The phonology interacts with the other components of the grammar. The stress differences between the noun and verb pairs discussed in the previous section (*subject* as noun or verb) are also predictable from the syntactic word category.

Intonation

> "What can *I* do, Tertius?" said Rosamond, turning her eyes on him again. That little speech of four words, like so many others in all languages, is capable by varied vocal inflexions of expressing all states of mind from helpless dimness to exhaustive argumentative perception, from the completest self-devoting fellowship to the most neutral aloofness.
>
> **GEORGE ELIOT,** *Middlemarch,* 1872

In chapter 6, we discussed pitch as a phonetic feature in reference to tone languages and intonation languages. In this chapter we have discussed the use of phonetic features to distinguish meaning. We can now see that pitch is a *phonemic* feature in tone languages such as Chinese, Thai, and Akan. We refer to these relative pitches as **contrasting tones**. In intonation languages such as English, pitch still plays an important role, but in the form of the **pitch contour** or **intonation** of the phrase or sentence.

In English, intonation may reflect syntactic or semantic differences. If we say *John is going* with a falling pitch at the end, it is a statement, but if the pitch rises at the end, it may be interpreted as a question. Similarly, *What's in the tea, honey?* may, depending on intonation, be a query to someone called "honey" regarding the contents of the tea (falling intonation on *honey*), or may be a query regarding whether the tea contains honey (rising intonation on *honey*).

A sentence that is ambiguous in writing may be unambiguous when spoken because of differences in the pitch contour, as we saw in the previous paragraph.

Here is a somewhat more subtle example. Written, sentence 1 is unclear as to whether Tristram intended for Isolde to read and follow directions, or merely to follow him:

1. Tristram left directions for Isolde to follow.

Spoken, if Tristram wanted Isolde to follow him, the sentence would be pronounced with a rise in pitch on the first syllable of *follow*, followed by a fall in pitch, as indicated (oversimplistically) in sentence 2.

Tristram left directions for Isolde to follow.

In this pronunciation of the sentence, the primary stress is on the word *follow*.

If the meaning is to read and follow a set of directions, the highest pitch comes on the second syllable of *directions*, as illustrated, again oversimplistically, in sentence 3.

Tristram left directions for Isolde to follow.

The primary stress in this pronunciation is on the word *directions*.

Pitch plays an important role in both tone languages and intonation languages, but in different ways, depending on the phonological system of the respective languages.

Sequential Constraints of Phonemes

> If you were to receive the following telegram, you would have no difficulty in correcting the "obvious" mistakes:
>
> BEST WISHES FOR VERY HAPPP BIRTFDAY
>
> because sequences such as BIRTFDAY do not occur in the language.
>
> **COLIN CHERRY,** *On Human Communication*, 1957

Suppose you were given the following four phonemes and asked to arrange them to form all *possible* English words:

/b/ /ɪ/ /k/ /l/

You would most likely produce the following:

/blɪk/
/klɪb/
/bɪlk/
/kɪlb/

These are the only permissible arrangements of these phonemes in English. */lbkɪ/, */ɪlbk/, */bkɪl/, and */ɪlkb/ are not possible English words. Although /blɪk/ and /klɪb/ are not now existing words, if you heard someone say:

"I just bought a beautiful new blick."

you might ask: "What's a blick?"

If, on the other hand, you heard someone say:

"I just bought a beautiful new bkli."

you might reply, "You just bought a new *what*?"

Your knowledge of English phonology includes information about what sequences of phonemes are permissible, and what sequences are not. After a consonant like /b/, /g/, /k/, or /p/, another stop consonant in the same syllable is not permitted by the phonology. If a word begins with an /l/ or an /r/, the next segment must be a vowel. That is why */lbɪk/ does not sound like an English word. It violates the restrictions on the sequencing of phonemes. People who like to work crossword puzzles are often more aware of these constraints than the ordinary speaker, whose knowledge, as we have emphasized, may not be conscious.

Other such constraints exist in English. If the initial sounds of *chill* or *Jill* begin a word, the next sound must be a vowel. The words /tʃʌt/ or /tʃon/ or /tʃæk/ are possible in English (*chut*, *chone*, *chack*), as are /dʒæl/ or /dʒil/ or /dʒalɪk/ (*jal*, *jeel*, *jolick*), but */tʃlɔt/ and */dʒpurz/ are not. No more than three sequential consonants can occur at the beginning of a word, and these three are restricted to /s/ + /p,t,k/ + /l,r,w,y/. There are even restrictions if this condition is met. For example, /stl/ is not a permitted sequence, so *stlick* is not a possible word in English, but *strick* is, along with *spew* /spju/, *sclaff* /sklæf/ (to strike the ground with a golf club), and *squat* /skwat/.

Other languages have different sequential restrictions. In Polish *zl* and *kt* are permissible syllable-initial combinations, as in /zlev/, "a sink," and /kto/, "who." Croatian permits words like the name *Mladen*. Japanese has severe constraints on what may begin a syllable; most combinations of consonants (e.g., /br/, /sp/) are impermissible.

The limitations on sequences of segments are called **phonotactic constraints**. Phonotactic constraints have as their basis the syllable, rather than the word. That is, only the clusters that can begin a syllable can begin a word, and only a cluster that can end a syllable can end a word.

In multisyllabic words, clusters that seem illegal may occur, for example the /kspl/ in *explicit* /ɛksplɪsɪt/. However, there is a syllable boundary between the /ks/ and /pl/, which we can make explicit using $: /ɛk $ splɪs $ ɪt/. Thus we have a permitted syllable coda /k/ that ends a syllable adjoined to a permitted onset /spl/ that begins a syllable. On the other hand, English speakers know that "condstluct" is not a possible word because the second syllable would have to start with an impermissible onset, either /stl/ or /tl/.

In Twi, a word may end only in a vowel or a nasal consonant. The sequence /pik/ is not a possible Twi word because it breaks the phonotactic rules of the language, whereas /mba/ is not a possible word in English, although it is a word in Twi.

All languages have constraints on the permitted sequences of phonemes, although different languages have different constraints. Just as spoken language has sequences of sounds that are not permitted in the language, so sign languages have forbidden combinations of features. For example, in the ASL compound for "blood" (red flow) discussed earlier, the total handshape must be assimilated, including the shape of the hand and the orientation of the fingers. Assimilation of just the handshape but not the finger orientation is impossible in ASL. The constraints may differ from one sign language to another, just as the constraints on sounds and sound sequences differ from one spoken language to another. A permissible sign in a Chinese sign language may not be a permissible sign in ASL, and vice versa. Children learn these constraints when they acquire the spoken or signed language, just as they learn what the phonemes are and how they are related to phonetic segments.

Lexical Gaps

The words *bot* [bat] and *crake* [kʰrek] are not known to all speakers of English, but they are words. On the other hand [bʊt] (rhymes with *put*), *creck* [kʰrɛk], *cruke* [kʰruk], *cruk* [kʰrʌk], and *crike* [kʰraɪk] are not now words in English, although they are possible words.

Advertising professionals often use possible but nonoccurring words for the names of new products. Although we would hardly expect a new product or company to come on the market with the name *Zhleet* [ʒlit]—an impossible word in English—we do not bat an eye at *Bic*, *Xerox* /ziraks/, *Kodak*, *Glaxo*, or *Spam* (a meat product, not junk mail), because those once nonoccurring words obey the phonotactic constraints of English.

A *possible word* contains phonemes in sequences that obey the phonotactic constraints of the language. An actual, occurring word is the union of a possible word with a meaning. Possible words without meaning are sometimes called nonsense words and are also referred to as **accidental gaps** in the lexicon, or **lexical gaps**. Thus "words" such as *creck* and *cruck* are nonsense words and represent accidental gaps in the lexicon of English.

Why Do Phonological Rules Exist?

> No rule is so general, which admits not some exception.
>
> **ROBERT BURTON,** *The Anatomy of Melancholy*, 1621

A very important question that we have not addressed thus far is: Why do grammars have phonological rules at all? In other words, why don't underlying or phonemic forms surface intact rather than undergoing various changes?

In the previous section we discussed *phonotactic constraints*, which are part of our knowledge of phonology. As we saw, phonotactic constraints specify which sound sequences are permissible in a particular language, so that in English *blick* is a possible word but **lbick* isn't. Many linguists believe that phonological rules exist to ensure that the surface or phonetic forms of words do not violate phonotactic constraints. If underlying forms remained unmodified, they would often violate the phonotactics of the language.

Consider, for example, the English past-tense rule and recall that it has two subrules. The first inserts a schwa when a regular verb ends in an alveolar stop (/t/ or /d/), as in *mated* [met**əd**]. The second devoices the past-tense morpheme /d/ when it occurs after a voiceless sound, as in *reaped* [rip**t**] or *peaked* [pʰik**t**].

Notice that the part of the rule that devoices /d/ reflects the constraint that English words may not end in a sequence consisting of a voiceless stop + d. Words such as [lɪpd] and [mɪkd] do not exist, nor could they exist. They are impossible words of English, just as [bkɪl] is.

More generally, there are no words that end in a sequence of obstruents whose voicing features do not match. Thus words such as [kasb], where the final two obstruents are [–voice] [+voice] are not possible, nor are words such as [kabs] whose final two obstruents are [+voice] [–voice]. On the other hand, [kasp] and [kɛbz] are judged to be possible words because the final two segments agree in voicing. Thus, there appears to be a general constraint in English, stated as follows:

(A) Obstruent sequences may not differ with respect to their voice feature at the end of a word.

We can see then that the devoicing part of the past-tense rule changes the underlying form of the past-tense morpheme to create a surface form that conforms to this general constraint.

Similarly, the schwa insertion part of the past-tense rule creates possible sound sequences from impossible ones. English does not generally permit sequences of sounds within a single syllable that are very similar to each other, such as [kk], [kg], [gk], [gg], [pp], [sz], [zs], and so on. (The words spelled *egg* and *puppy* are phonetically [ɛg] and [pʌpɪ].) Thus the schwa insertion rule separates sequences of sounds that are otherwise not permitted in the language because they are too similar to each other, for example, the sequence of /d/ and /d/ in /mɛnd + d/, which becomes [mɛ̃ndəd] *mended*, or /t/ and /d/ in /part + d/, which becomes [$\mathrm{p^h}$artəd] *parted*. The relevant constraint is stated as follows:

(B) Sequences of obstruents that differ at most with respect to voicing are not permitted within English words.

Constraints such as (A) and (B) are far more general than particular rules like the past-tense rule. For example, constraint B might also explain why an adjective such as *smooth* turns into the abstract noun *smoothness*, rather than taking the affix *-th* [θ], as in *wide-width*, *broad-breadth*, and *deep-depth*. Suffixing *smooth* with *-th* would result in a sequence of too similar obstruents, smoo[ðθ], which differ only in their voicing feature. This suggests that languages may satisfy constraints in various grammatical situations.

Thus, phonological rules exist because languages have general principles that constrain possible sequences of sounds. The rules specify minimal modifications of the underlying forms that bring them in line with the surface constraints. Therefore, we find different variants of a particular underlying form depending on the phonological context.

It has also been proposed that a universal set of phonological constraints exists, and that this set is ordered, with some constraints being more highly ranked than others. The higher the constraint is ranked, the more influence it exerts on the language. This proposal, known as **Optimality Theory**, also holds that the particular constraint rankings can differ from language to language, and that the different rankings generate the different sound patterns shown across languages. For example, constraint B is highly ranked in English; and so we have the English past-tense rule, as well as many other rules, including the plural rule (with some modification), that modify sequences of sounds that are too similar. Constraint B is also highly ranked in other languages, for example, Modern Hebrew, in which suffixes that begin with /t/ are always separated from stems ending in /t/ or /d/ by inserting [e], as in /kiʃat + ti/ → [kiʃatetɪ] meaning "I decorated." In Berber, similar consonants such as tt, dd, ss, and so on can surface at the end of words. In this language, constraint B is not highly ranked; other constraints outrank it and therefore exert a stronger effect on the language, notably constraints that require that surface forms not deviate from corresponding underlying forms. These constraints, known as *faithfulness constraints*, compete in the rankings with constraints that modify the underlying forms. Faithfulness constraints reflect the drive among languages to want a morpheme to have a single identifiable form, a drive that is in competition with constraints such as A and B. In the case of the English past-tense morpheme, the drive toward a single morpheme shows up in the spelling, which is always *-ed*.

In our discussion of syntactic rules in chapter 4, we noted that there are principles of Universal Grammar (UG) operating in the syntax. Two examples of this are the principle that transformational rules are structure dependent and the constraint that movement rules may not move phrases out of coordinate structures. If Optimality Theory is correct, and universal phonological constraints exist that differ among languages only in their rankings, then phonological rules, like syntactic rules, are constrained by universal principles. The differences in constraint rankings across languages are in some ways parallel to the different parameter settings that exist in the syntax of different languages, also discussed in chapter 4. We noted that in acquiring the syntax of her language, the young child must set the parameters of UG at the values that are correct for the language of the environment. Similarly, in acquiring the phonology of her language, the child must determine the correct constraint rankings as evidenced in the input language. We will have more to say about language acquisition in chapter 8.

Phonological Analysis

> Out of clutter, find simplicity.
> From discord, find harmony.
>
> **ALBERT EINSTEIN** (1879–1955)

Children recognize phonemes at an early age without being taught, as we shall see in chapter 8. Before reading this book, or learning anything about phonology, you knew a *p* sound was a phoneme in English because it contrasts words like *pat* and *cat*, *pat* and *sat*, *pat* and *mat*. But you probably did not know that the *p* in *pat* and the *p* in *spit* are different sounds. There is only one /p/ phoneme in English, but that phoneme has more than one allophone, including an aspirated one and an unaspirated one.

If a non-English-speaking linguist analyzed English, how could this fact about the sound *p* be discovered? More generally, how do linguists discover the phonological system of a language?

To do a phonological analysis, the words to be analyzed must be transcribed in great phonetic detail, because we do not know in advance which phonetic features are distinctive and which are not.

Consider the following Finnish words:

1.	[kudot]	"failures"	**5.**	[madon]	"of a worm"
2.	[kate]	"cover"	**6.**	[maton]	"of a rug"
3.	[katot]	"roofs"	**7.**	[ratas]	"wheel"
4.	[kade]	"envious"	**8.**	[radon]	"of a track"

Given these words, do the voiceless/voiced alveolar stops [t] and [d] represent different phonemes, or are they allophones of the same phone?

Here are a few hints as to how a phonologist might proceed:

1. Check to see if there are any minimal pairs.
2. Items (2) and (4) are minimal pairs: [kate] "cover" and [kade] "envious." Items (5) and (6) are minimal pairs: [madon] "of a worm" and [maton] "of a rug."
3. [t] and [d] in Finnish thus represent the distinct phonemes /t/ and /d/.

That was an easy problem. Now consider the following data from English, again focusing on [t] and [d] together with the alveolar flap [ɾ] and primary stress ´:

[ráɪt]	"write"	[ráɪɾər]	"writer"
[déɾə]	"data"	[dét]	"date"
[mǽd]	"mad"	[mǽt]	"mat"
[bətróð]	"betroth"	[lǽɾər]	"ladder"
[lǽɾər]	"latter"	[dístə̃ns]	"distance"
[ráɪɾər]	"rider"	[ráɪd]	"ride"
[déɾɪŋ]	"dating"	[bɛ́dsaɪd]	"bedside"
[mʌ́ɾər]	"mutter"	[túɾər]	"tutor"
[mǽɾər]	"madder"	[mǽdnɪs]	"madness"

A broad examination of the data reveals minimal pairs involving [t] and [d], so clearly /t/ and /d/ are phonemes. We also see some interesting homophones, such as *ladder* and *latter*, and *writer* and *rider*. And the flap [ɾ]? Is it a phoneme? Or is it predictable somehow? At this point the linguist undertakes the tedious task of identifying *all* of the immediate environments for [t], [d], and [ɾ], using # for a word boundary:

[t]: áɪ_#, é_#, ǽ_#, ə_r, s_ə, #_ú
[d]: #_é (3 times), ǽ_#, #_í, áɪ_#, ɛ́_s, ǽ_n
[ɾ]: áɪ_ə (2 times), é_ə, ǽ_ə (3 times), é_ɪ, ú_ə, ʌ́_ə

It does not appear at this point that anything systematic is going on with vowel or consonant quality, so we abstract the data a little, using *v* for an unstressed vowel, *v́* for a stressed vowel, C for a consonant, and # for a word boundary:

[t]: v́_#, #_v́, C_v, v_C
[d]: #_v́, v́_#, v́_C
[ɾ]: v́_v

Now we see clearly that [ɾ] is in complementary distribution with both [t] and [d]. It occurs only when preceded by a stressed vowel and followed by an unstressed vowel, and neither [t] nor [d] ever do. We may conclude, based on these data, that [ɾ] is an allophone of both /t/ and /d/. We tentatively propose the "alveolar flap rule":

> An alveolar stop becomes a flap in the environment between a stressed and unstressed vowel.

The phonemic forms lack a flap, so that *writer* is phonemically /raɪtər/ and *rider* is /raɪdər/, based on [raɪt] and [raɪd]. Similarly, we can propose /mædər/ for *madder* based on [mæd] and [mædnɪs], and /detɪŋ/ for *dating* based on [det]. But we don't have enough information to determine phonemic forms of *data*, *latter*, *ladder*, *tatter*, and *tutor*. This is typically the case in actual analyses. Rarely is there sufficient evidence to provide all the answers.

Finally, consider these data from Greek, focusing on the following sounds:

[x]	voiceless velar fricative
[k]	voiceless velar stop
[c]	voiceless palatal stop
[ç]	voiceless palatal fricative

1.	[kano]	"do"	**9.**	[çeri]	"hand"
2.	[xano]	"lose"	**10.**	[kori]	"daughter"
3.	[çino]	"pour"	**11.**	[xori]	"dances"
4.	[cino]	"move"	**12.**	[xrima]	"money"
5.	[kali]	"charms"	**13.**	[krima]	"shame"
6.	[xali]	"plight"	**14.**	[xufta]	"handful"
7.	[çeli]	"eel"	**15.**	[kufeta]	"bonbons"
8.	[ceri]	"candle"	**16.**	[oçi]	"no"

To determine the status of [x], [k], [c], and [ç], you should answer the following questions.

1. Are there are any minimal pairs in which these sounds contrast?
2. Are any noncontrastive sounds in complementary distribution?
3. If noncontrasting phones are found, what are the phonemes and their allophones?
4. What are the phonological rules by which the allophones can be derived?

1. By analyzing the data, we find that [k] and [x] contrast in a number of minimal pairs, for example, in [kano] and [xano]. [k] and [x] are therefore distinctive. [c] and [ç] also contrast in [çino] and [cino] and are therefore distinctive. But what about the velar fricative [x] and the palatal fricative [ç]? And the velar stop [k] and the palatal stop [c]? We can find no minimal pairs that would conclusively show that these represent separate phonemes.

2. We now proceed to answer the second question: Are these noncontrasting phones, namely [x]/[ç] and [k]/[c], in complementary distribution? One way to see if sounds are in complementary distribution is to list each phone with the environment in which it is found, as follows:

Phone	**Environment**
[k]	before [a], [o], [u], [r]
[x]	before [a], [o], [u], [r]
[c]	before [i], [e]
[ç]	before [i], [e]

We see that [k] and [x] are not in complementary distribution; they both occur before back vowels. Nor are [c] and [ç] in complementary distribution. They both occur before front vowels. But the stops [k] and [c] are in complementary distribution; [k] occurs before back vowels and [r], and never occurs before front vowels. Similarly, [c] occurs only before front vowels and never before back vowels or [r]. Finally, [x] and [ç] are in complementary distribution for the same reason. We therefore conclude that [k] and [c] are allophones of one phoneme, and the fricatives [x] and [ç] are also allophones of one phoneme. The pairs of allophones also fulfill the criterion of phonetic similarity. The first two are [–anterior] stops; the second are [–anterior] fricatives. (This similarity discourages us from pairing [k] with [ç], and [c] with [x], which are less similar to each other.)

3. Which of the phone pairs are more basic, and hence the ones whose features would define the phoneme? When two allophones can be derived from one phoneme, one selects as the underlying segment the allophone that makes the rules and the phonemic feature matrix as simple as possible, as we illustrated with the English unaspirated and aspirated voiceless stops.

In the case of the velar and palatal stops and fricatives in Greek, the rules appear to be equally simple. However, in addition to the simplicity criterion, we wish to state rules that have natural phonetic explanations. Often these turn out to be the simplest solution. In many languages, velar sounds become palatal before front vowels. This is an assimilation rule; palatal sounds are produced toward the front of the mouth, as are front vowels. Thus we select /k/ as a phoneme with the allophones [k] and [c], and /x/ as a phoneme with the allophones [x] and [ç].

4. We can now state the rule by which the palatals can be derived from the velars.

Palatalize velar consonants before front vowels.

Using feature notation we can state the rule as:

[+velar] → [+palatal] / ____ [–back]

Because only consonants are marked for the feature [velar], and only vowels for the feature [back], it is not necessary to include the features [consonantal] or [syllabic] in the rule. We also do not need to include any other features that are redundant in defining the segments to which the rule applies or the environment in which the rule applies. Thus [+palatal] in the change part of the rule is sufficient, and the feature [–back] also suffices to specify the front vowels. The simplicity criterion constrains us to state the rule as simply as we can.

Summary

Part of one's knowledge of a language is knowledge of the **phonology** or sound system of that language. It includes the inventory of **phones—**which are the phonetic sounds that occur in the language—and the ways in which they pattern. This patterning determines the inventory of **phonemes—**the abstract basic units that differentiate words.

When similar phones occur in **complementary distribution**, they are **allophones—**predictable phonetic variants—of one phoneme. Thus the aspirated [p^h] and the unaspirated [p] are allophones of the phoneme /p/ because they occur in different phonetic environments.

Some phones may be allophones of more than one phoneme. There is no one-to-one correspondence between the phonemes of a language and their allophones. In English, for example, stressed vowels become unstressed according to regular rules, and ultimately reduce to schwa [ə], which is an allophone of each English vowel.

Phonological segments—phonemes and phones—are composed of **phonetic features** such as *voiced*, *nasal*, *labial*, and *continuant*, whose presence or absence is indicated by + or – signs. *Voiced*, *continuant*, and many others are **distinctive features—**they can contrast words. Other features like *aspiration* are **nondistinctive** and are predictable from phonetic context. Some features like *nasal* may be distinctive for one class of sounds (e.g., consonants) but nondistinctive for a different class of sounds (e.g., vowels). Phonetic features that are nondistinctive in one language may be distinctive in another. Aspiration is distinctive in Thai and nondistinctive in English.

When two distinct words are distinguished by a single phone occurring in the same position, they constitute a **minimal pair**, e.g., *fine* [faɪn] and *vine* [vaɪn]. Minimal pairs also occur in sign languages. Signs may contrast by handshape, location, and movement.

Words in some languages may also be phonemically distinguished by **prosodic** or **suprasegmental** features, such as pitch, stress, and segment length. Languages in which syllables or words are contrasted by pitch are called **tone languages**. **Intonation** languages may use pitch variations to distinguish meanings of phrases and sentences.

The relationship between phonemic representation and phonetic representation (pronunciation) is determined by phonological rules. Phonological rules apply to phonemic strings and alter them in various ways to derive their phonetic pronunciation, or in the case of signed languages, their hand configuration. They may be **assimilation rules**, **dissimilation rules**, rules that *add* **nondistinctive features**, **epenthetic** rules that insert segments, **deletion** rules, and **metathesis** rules that reorder segments.

Phonological rules generally refer to entire classes of sound. These are **natural classes**, characterized by a small set of phonetic features shared by all the members of the class, e.g., [–continuant], [–voiced], to designate the natural class of voiceless stops.

Linguists may use a mathematical-like formulation to express phonological rules in a concise way. For example, the rule that nasalizes vowels when they occur before a nasal consonant may be written V → [+nasal] / __ [+nasal].

Morphophonemic rules apply to specific morphemes, e.g., the plural morpheme /z/ is phonetically [z], [s], or [əz], depending on the final phoneme of the noun to which it is attached.

The phonology of a language also includes sequential constraints (**phonotactics**) that determine which sounds may be adjacent within the syllable. These determine what words are possible in a language, and what phonetic strings are impermissible. Possible but nonoccurring words constitute **accidental gaps** and are **nonsense words**, e.g., *blick* [blɪk].

Phonological rules exist in part to enforce phonotactic constraints. **Optimality Theory** hypothesizes a set of ranked constraints that govern the phonological rules.

To discover the phonemes of a language, linguists (or students of linguistics) can use a methodology such as looking for minimal pairs of words, or for sounds that are in complementary distribution.

The phonological rules in a language show that the phonemic shape of words is not identical with their phonetic form. The phonemes are not the actual phonetic sounds, but are abstract mental constructs that are realized as sounds by the operation of rules such as those described in this chapter. No one is taught these rules, yet everyone knows them subconsciously.

References for Further Reading

Anderson, S. R. 1985. *Phonology in the twentieth century: Theories of rules and theories of representations*. Chicago: University of Chicago Press.

Bybee, J. 2002. *Phonology and language use*. Cambridge, UK: Cambridge University Press.

Chomsky, N., and M. Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

Clements, G. N., and S. J. Keyser. 1983. *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.

Goldsmith, J. A. (ed.). 1995. *The handbook of phonological theory*. Cambridge, MA: Blackwell.

Gussman, E., S. R. Anderson, J. Bresnan, B. Comrie, W. Dressler, and C. J. Ewan. 2002. *Phonology: Analysis and theory*. Cambridge, UK: Cambridge University Press.

Hogg, R., and C. B. McCully. 1987. *Metrical phonology: A coursebook*. Cambridge, UK: Cambridge University Press.

Hyman, L. M. 1975. *Phonology: Theory and analysis*. New York: Holt, Rinehart & Winston.

Kaye, Jonathan. 1989. *Phonology: A cognitive view*. Hillsdale, NJ: Erlbaum.

Kenstowicz, M. J. 1994. *Phonology in generative grammar*. Oxford, UK: Blackwell Publications.