FILE 3.1

Phonotactic Constraints and Foreign Accents

3.1.1 Phonotactic Constraints

In Chapter 2, we discussed the sounds of a language; part of knowing a language is knowing what sounds are part of that language and how to pronounce them. In addition to this, however, we also need to know how those sounds work together in a system: where sounds can occur and how they influence one another.

In every language there are restrictions on the kinds of sounds and sound sequences possible in different positions in words (particularly at the beginning and end). These restrictions can be formulated in terms of rules stating which sound sequences are possible in a language and which are not. Restrictions on possible combinations of sounds are known as **phonotactic constraints.**

We can investigate examples of restrictions on consonant sequences in more detail by considering some in a language we know very well—English. To start with, any consonant of English may occur initially (at the beginning) in words except for two: [ʒ] and [ŋ]. While some speakers do pronounce these sounds in borrowed words such as *Jacques* and *Nguyen,* no native English word begins with them. A large number of two-consonant combinations also occur word-initially, with a stop or fricative being followed by a liquid or glide:

(1)	[bɹ]	*bring*	[gl]	*glean*	[mj]	*music*	[kw]	*quick*
	[θɹ]	*three*	[fl]	*fly*	[hj]	*humor*	[sw]	*sweet*

In addition, [s] can also be followed by voiceless and nasal stops (as in *stay, small*) and by [f] and [v] in a small number of borrowed words (*sphere, svelte,* etc.). [ʃ] can be followed by a nasal stop or a liquid, but only [ʃɹ] is a cluster native to English (e.g., *shrink*). The others are present only in borrowings from Yiddish and German (*Schlemiel* 'clumsy person,' *Schnook,* 'fool,' *Schwinn*).

In addition to having constraints concerning which particular sounds are permitted to occur together or in certain positions, languages have phonotactic constraints about what kinds of sounds are permitted to occur in different positions. Many of these constraints regard syllable types (see File 2.1 for syllable structure).

Languages generally prefer syllables made up of a consonant (C) first, and a vowel (V) second, but some languages allow a syllable to begin with more than one consonant. For instance, English allows up to three consonants to start a word, provided the first is [s], the second [p], [t], or [k], and the third [l], [ɹ], [j], or [w] (see below). There is a wide variety of syllable types in English, as illustrated in (2).

(2)	V	*a*	CV	*no*	CCV	*flew*	CCCV	*spree*
	VC	*at*	CVC	*not*	CCVC	*flute*	CCCVC	*spleen*
	VCC	*ask*	CVCC	*ramp*	CCVCC	*flutes*	CCCVCC	*strength*
	VCCC	*asked*	CVCCC	*ramps*	CCVCCC	*crafts*	CCCVCCC	*strengths*

Other languages, however, do not have such a large number of syllable structures, as the lists in (3) illustrate. (Hebrew CVCC syllables are allowed only at the end of a word, and only if the final consonant is [t].)

(3)

Hawaiian	Indonesian	Hebrew
CV	CV	CV
V	V	CCV
	VC	CCVC
	CVC	CVC
		CVCC

Notice that this means that Indonesian has clusters only in the middle of words; that is, there are no clusters initially or finally. Hawaiian does not permit clusters in any position. Meanwhile, even though Hebrew permits both initial and final clusters, it does not allow a single vowel to be a syllable by itself. Every language has its own set of permitted segmental sequences.

The phonotactic constraints of a language will generally apply to every word in the language, native or not. Therefore, languages seek to overcome problems of borrowing a foreign word that violates their phonotactics. For instance, in English, two stops cannot come at the beginning of words, nor can stop plus nasal combinations. So, in order to pronounce the borrowed words *Ptolemy* and *gnostic* more easily, English speakers simply drop the first consonant and pronounce the words [tɑləmi] and [nɑstɪk], respectively. Or, speakers may insert a vowel between the two consonants, as in the pronunciation of the words *Gdansk* and *knish* as [gədænsk] and [kənɪʃ]. Both of these alterations eliminate violations of the phonotactic constraints of English.

As these examples from English illustrate, there are different ways of handling phonotactic problems. Japanese and Finnish provide us with additional examples. Japanese and Finnish generally avoid syllables containing sequences of consonants. When borrowing a foreign word that violates the language's syllable structure, the two languages must change it somehow to fit. There are two ways that borrowed words with consonant clusters are "repaired." One is to drop or delete one of the consonants; the other is to insert a vowel to separate the consonants. Finnish opts for deletion. In loan words, Finnish drops the first of a series of consonants that do not conform to its phonotactics. Thus, Germanic *Strand* (CCCVNC) ends up as *ranta* 'beach' (CVNCV) in Finnish, and *glass* becomes *lasi*. Note also the addition of a final vowel to avoid a consonant in syllable-final position.

The other way to break up consonant clusters is used in Japanese. Japanese inserts vowels into the cluster, so that, for example, a CCC sequence will end up as CVCVCV. The vowel insertion is predictable and rule-governed: the vowel [o] is inserted after [t] and [d], while the vowel [ɯ] is inserted after all other consonants. Thus, we can predict the form of new words in Japanese that have been borrowed from English. For example, when the English word *birth control* was borrowed into Japanese, it became [ba:sɯ kontoro:rɯ]. Note that the nasals [n] and [m] are allowed to occur syllable-finally in Japanese, although no other consonants are.

/bəɹθ/ → [ba:sɯ]
/kəntɹol/ → [kontoro:rɯ]

[ɯ] is inserted the end of [ba:sɯ] and [kontoro:rɯ] to keep the word-final syllables from ending in a consonant. The second [o] in [kontoro:rɯ] is inserted to prevent [t] and [r] from

forming a consonant cluster. Notice also that Japanese substitutes other sounds for some of the English sounds. This will be discussed in 3.1.3.

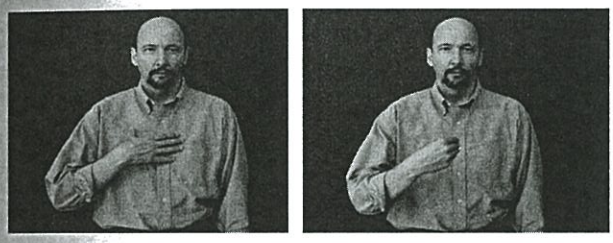
3.1.2 Phonotactic Constraints in Signed Languages

There are similar kinds of constraints on what sorts of segment combinations are and are not allowed in various signed languages. As with the phonotactic constraints for syllable structures and for consonants and vowels in spoken languages described above, constraints on syllable structure and on what sorts of handshapes and movements can appear adjacent to one another in signed languages differ from language to language. The phonotactic constraints discussed in this section are specific to ASL.

First we will consider restrictions on syllable structure; there will be two examples.[1] It was mentioned above that in Hebrew, a vowel alone cannot serve as a syllable: there is a minimum requirement that a syllable in Hebrew contain at least two segments. There is a similar minimum requirement for ASL syllables: a monosyllabic sign cannot consist of just one handshape, one location, and one orientation; at least one of these elements is required to change in order to form a grammatical syllable.

The second example we will consider is when changes of handshape are allowed. Many signs include a change of handshape during movement of the hands from one location to another; other signs involve handshape changes that occur while the hands are held stationary at some particular place. In ASL, handshape changes may always occur during movement. The sign WHITE, shown in (4), provides a good example of this.

(4) ASL: WHITE



Likewise, if a sign comprises only one place of articulation without movement to another place of articulation, handshape change can occur while the hand is kept at that location. However, if a sign involves the hands being at some place and then moving, or moving and then winding up at some place, then the change of handshape must take place during the movement. It is not permitted in ASL for handshape to change while the hands are held at some particular location if there is a movement component of that sign.

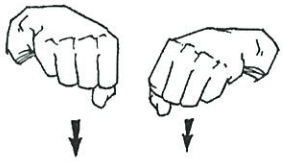
Not only do signed languages have syllable structure constraints, but also there are constraints on which segments can be adjacent to one another in the same way that certain sound combinations are not allowed in some languages (like word-initial /fk/ in English). For example, in any given signed language, there may be certain handshapes which—though parts of the system of handshapes in that language—are not allowed by the grammar to appear adjacent to one another within a word.

---

[1] It probably seems very peculiar to think about signs as having syllables! Nonetheless, signs can be broken down into prosodic units just like spoken words. If you are not a signer, though, it is very difficult to figure out what might comprise a signed syllable. (We also lack intuitions about syllable structure of spoken languages that we do not know.) Therefore, we will simply take it as an underlying assumption that signed languages have syllables, and go from there.

An interesting phonotactic constraint that does not have any obvious direct parallel in spoken languages but that seems fairly uniform among signed languages pertains to the fact that in signed languages there are two possible primary articulators, namely, the right hand and the left hand. In all signed languages studied to date, a signer may be right-hand dominant or left-hand dominant (which roughly corresponds to whether the signer is right- or left-handed). The dominant hand is the one that the signer will use to perform all one-handed signs. (If a right-handed signer were to injure her right arm or happened to be carrying a large box under her right arm, she might temporarily switch and use her left hand to sign with, but no signer switches back and forth between hands as a matter of course: this would be considered extremely aberrant to any native speaker of a signed language.) Interestingly, there are very specific restrictions on what the non-dominant hand may do in any given two-handed sign. If both hands are moving, then the non-dominant hand must have the same handshape, orientation, and motion as the dominant hand; that is, in signs where both hands are moving, there is a symmetry constraint. It is important to note that although the motion must be the same, the direction of the motion may be parallel or anti-parallel. An example of an ASL sign that follows the symmetry constraint is CAN, meaning 'be able to,' illustrated in (5). Although the non-dominant hand does move in this sign, it mirrors exactly the shape, orientation, and movement of the dominant hand.

(5) ASL: CAN



The non-dominant hand may also participate in a sign by remaining stationary while the dominant hand moves. This is exemplified in the ASL sign CHOCOLATE in (6): the two hands have different handshapes, but the non-dominant hand (the lower hand, which is held flat) is not moving.

(6) ASL: CHOCOLATE

A grammatical sign in any signed language cannot have both hands moving unless they both have the same handshape and orientation and are performing the same kind of movement. Interestingly, in Signed Mandarin (which is not a signed language, but rather a signed code for Mandarin Chinese; see File 1.5) there are certain signs that have been introduced by hearing (non-native signer) instructors at deaf schools that do not follow this rule. The fact that the Signed Mandarin words do not follow the universal rule for signed languages is yet more evidence that signed codes and signed languages differ! For example, in the signed Mandarin sign for 'ink,' both hands are moving, so the sign should follow the symmetry constraint. However, in this sign the dominant hand is facing toward the signer, and the non-dominant hand away from the signer, so they have different orientations; the dominant hand moves in a path away from the body while the non-dominant hand moves

from side to side, so they have different movement; and the dominant hand has one finger extended while the non-dominant hand has three fingers extended, so they have different handshapes. Thus this sign is ungrammatical for three reasons.

There has been an attempt (again by non-native signers) to introduce such signs from Signed Mandarin into Taiwan Sign Language, but the reaction among the native signers is that these signs are not possible in their language. It is exactly as though somebody told you that /kpflus/ was a new word of English; you wouldn't accept it!

3.1.3 Foreign Accents

Applying the phonotactic constraints of one language while speaking another is an important source of foreign accents. A Spanish speaker, for example, may pronounce *student* as [ɛstudɛnt] because in Spanish, the consonant clusters [st], [sk], and [sp] are not permitted to occur at the beginning of a word without being preceded by a vowel—as in the words *estudiante* 'student,' *escuela* 'school,' and *espalda* 'shoulder.' The Spanish speaker who says [ɛstudɛnt] is simply applying the phonotactic constraints of Spanish when speaking English words.

Another source of foreign accents is the fact that not all sound systems are the same. In File 2.4, we discussed many sounds that are speech sounds in other languages but that English does not use. Similarly, there are sounds that are part of the sound inventory of English that are not used by other languages, and we can detect this when we hear non-native speakers of English pronounce English words. For example, French speakers often pronounce English *this* [ðɪs] as [zɪs] and *thin* [θɪn] as [sɪn]. The reason for this difference is that the phonemic inventory of French does not contain [ð] or [θ]. Because of this, French speakers substitute the nearest sounds available in their phonemic inventory: in this case, [z] and [s]. This process is called **sound substitution,** by which sounds that exist in a language a speaker knows are used to replace sounds that do not exist in that language when pronouncing the words of a foreign language.

Notice that the sounds that French speakers use to replace [ð] and [θ] are very similar to the sounds they replace. [ð] is replaced by [z], another voiced fricative that is pronounced at a similar place of articulation, and /θ/ is replaced by the corresponding voiceless fricative [s]. In most cases, speakers will replace a sound that does not exist in their language by the sound in their native language that is most similar.

Another familiar example involves the pronunciation of German by some speakers of English. German has a voiceless velar fricative [x]. The phonemic inventory of English, of course, lacks this sound, though we do have a voiceless velar stop [k]. Most speakers of English substitute [k] for [x] in a German word like *Bach* [bɑx], producing [bɑk] instead. Another example of the same substitution is the way most American English speakers pronounce *Lebkuchen* [lebkuxən] 'Christmas cookie' as [leɪbkukən]. Some English speakers, striving for a more "German-like" pronunciation, will pronounce it instead as [leɪbkuhən]. Why do you suppose an English speaker might substitute [h] for [x]?

FILE 3.2

Phonemes and Allophones

3.2.1 Predicting the Occurrence of Sounds

As discussed in File 3.1, there are restrictions on which sounds can occur in certain environments. In some cases, the environment allows us to predict the occurrence of the sounds. Consider the sounds [k] and [g]. In both Kikamba (a Bantu language spoken in Kenya) and English, we can hear the sounds [k] and [g]. The Kikamba word [kosuuŋga] 'to guard' contains both phones, as does the English word [kɑgneɪt] *cognate*. The difference between Kikamba and English lies in the way the two sounds contribute to the identity of a word. In English, the two phones can distinguish words, as shown by words like [tæk] *tack* and [tæg] *tag*, where the difference between [k] and [g] affects which word is produced: if a [k] is pronounced, the word is *tack*, while a [g] in the same place gives a different word, *tag*.

In this sense, phonologists say that the occurrence of these two sounds in English is unpredictable, since we cannot look at the rest of the word and determine which sound will occur. That is, if we know that a word in English begins with [tæ], we cannot predict whether the word will end with [k] or [g] since both *tack* and *tag* are different, but possible, words.

In Kikamba, on the other hand, the sounds [k] and [g] *are* predictable from their environment. Sounds are predictable when we expect to see one sound but not the other based upon the sounds that precede and/or follow it. In Kikamba, the only consonant that can come directly after an [ŋ] is [g], and [g] can only come immediately after [ŋ]. The combination [ŋk] does not occur in Kikamba (see Roberts-Kohno 2000). So, if there is a velar stop in a word in Kikamba, we can predict whether it will be a [k] or a [g]: it will be a [g] if it is immediately preceded by [ŋ]; otherwise, it will be a [k]. However, in English we cannot make this prediction: the sound [k] does appear after the sound [ŋ], as in [æŋkɪ] *anchor*, as does the sound [g], as in [æŋgɪ] *anger*. In Kikamba, then, the distribution of the sounds [k] and [g] are predictable, while in English, it is not.

To illustrate how strong this distribution is in Kikamba, consider the case where you have a word with a [k] such as *katala*, the base form from which conditional forms of the verb 'to count' are built. To say 'if you count,' you add an [o] to the front of the word: [okatala]. But to say 'if I count,' you add an [ŋ]. Even though this word has a [k] in it, we have seen that it is a rule in this language that [k] cannot appear after [ŋ]—so the [k] appears as a [g] instead: [ŋgatala] 'if I count.' This type of alternation does not happen in English, because both [k] and [g] can occur after [ŋ].

So while Kikamba and English both use the phones [k] and [g], the languages differ in that in Kikamba we can predict the occurrence of one versus the other, and in English we cannot. If someone learning Kikamba were to use [k] after [ŋ], the meaning of the word would not change. Instead, a native speaker of Kikamba might think that the speaker sounded funny, had an accent, or had mispronounced the word. On the other hand, if a learner of English were to make the same substitution in English, pronouncing [k] after [ŋ] in *anger*, the word would change. A native speaker of English would hear *anchor*, rather than *anger*.

3.2.2 Allophones and Phonemes

In every language, certain sounds pattern together as if they were simply variants of the "same" sound, instead of different sounds that can be used to distinguish words, even though they may be phonetically distinct. For example, the sounds [k] and [g] are clearly different sounds: we use them to make the contrast between different words in English, as we saw in Section 3.2.1. But as we also saw, these two sounds are completely predictable in Kikamba. In Kikamba, then, these sounds can be thought of as variants of the "same" sound, because in any given context, if there is some velar stop sound, we can predict which one ([k] or [g]) it will be.

Similarly, if you ask a native speaker of English how many different sounds are represented by the underlined letters in the words *pin, bin,* and *spin,* they will probably say "two," grouping the aspirated [pʰ] of *pin* and unaspirated [p] of *spin* together. Though [pʰ] and [p] are phonetically different sounds, native English speakers often overlook this difference and may even consider them to be the "same" sound.

One of the goals of this file is to help you understand more clearly the distinction between "same" and "different" sounds. To do this, we will discuss the terms **allophone** and **phoneme.** Since these concepts are the crux of phonological analysis, it is important that they be clearly understood. Perhaps the best way to start to explain these terms is through examples. On a separate piece of paper, transcribe the following five words in IPA:

(1) top stop little kitten hunter

It is likely that you transcribed all of these words with a [t], like the following:

(2) [tɑp] [stɑp] [lɪtl̩] [kɪtn̩] [hʌntɻ̩]

This is good, because the [t] in each of these words is, for most speakers, the "same" sound as in the others. But, in fact, the physical reality (the acoustic phonetic fact) is that the 't' you transcribed in those five examples is pronounced slightly differently from one example to the next. To illustrate this, pronounce the five words again. Concentrate on what the 't' sounds like in each example, but be sure to say them as you normally would if you were talking to a friend—that is, don't try to enunciate them abnormally clearly.

What differences did you notice? Compare, for example, the 't' of *top* to that of *stop.* You should be able to detect a short burst or puff of air after the 't' in *top* that is absent in *stop.* That puff of air is what we have called aspiration (see Section 2.6.5), which is transcribed with a superscripted [ʰ]. So while a native speaker might think of the 't' sound in *top* and *stop* as being the same sound, the 't' is actually pronounced differently in each word. This difference can be captured in the transcription, as in [tʰɑp] and [stɑp], respectively.

Now say the words *little* and *kitten.* We might say that the 't' in *little* sounds "softer" than the one in *stop,* and is clearly voiced. For most speakers of American English (but not of British English), the 't' in words like *little* is pronounced as a flap, [ɾ], much like the *r* in Spanish in words like [paɾa] 'for' and [toɾo] 'bull' (see Section 2.2.5). English *kitten,* on the other hand, is pronounced with the same sound we hear in the expression *uh-oh,* a glottal stop [ʔ]. So, we could transcribe *little* and *kitten* as [lɪɾl̩] and [kɪʔn̩], respectively.

For some speakers of American English, in casual speech words like *hunter* are pronounced with no 't' at all, but rather as [hʌnɻ̩]. Try to say it this way and see if it sounds like something you've heard before. In any case, while you may have initially transcribed the five words above with a [t], they may also be transcribed in a way that reflects the different pronunciations of that sound, as in the following:

(3) [tʰɑp] [stɑp] [lɪɾl̩] [kʰɪʔn̩] [hʌnɻ̩]

To a native speaker, all of the words above may seem to have a 't' in them, at least at some psychological level.[1] Evidence of this lies in the fact that we may transcribe them with a 't,' at least until trained in transcription. Someone who lacks linguistic training would probably not hesitate to state that all the above words have a 't' and would need to be convinced that subtle differences, like aspiration, exist among them. In this sense, the above words do have a 't.' On the other hand, we can observe that the 't' may be pronounced in several different ways.

Unlike a speaker of English, a native speaker of Hindi does not ignore the difference between aspirated and unaspirated sounds when speaking or hearing Hindi. To a speaker of Hindi, the aspirated sound [pʰ] is as different from unaspirated [p] as [pʰ] is from [b] to our ears. The difference between aspirated and unaspirated stops must be noticed by Hindi speakers because their language contains many words that are pronounced in nearly the same way, except that one word will have an aspirated stop where the other has an unaspirated stop. The data in (4) illustrate this.

(4)

Hindi	**Gloss**
[pʰəl]	'fruit'
[pəl]	'moment'
[bəl]	'strength'

A native speaker of English may not be aware of the difference between aspirated and unaspirated stops because aspiration will never make a difference in the meanings of English words. If we hear someone say [mæp] and [mæpʰ], we may recognize them as different pronunciations of the same word *map,* but not as different words. Because of the different ways in which [p] and [pʰ] affect meaning distinctions in English and Hindi, these sounds have different values in the phonological systems of the two languages. We say that these two sounds are **noncontrastive** in English, because interchanging the two does not result in a change of meaning. In Hindi, on the other hand, [p] and [pʰ] are **contrastive** because replacing one sound with the other in a word can change the word's meaning. The meaning of a word with [p] (e.g. 'moment') contrasts with the meanings of a similar word with [pʰ] (e.g. 'fruit'). We will have more to say about this terminological distinction below.

Linguists attempt to characterize these different relations between sounds in language by grouping the sounds in a language's sound inventory into classes. Each class contains all of the sounds that a native speaker considers as the "same" sound. For example, [t] and [tʰ] in English would be members of the same class. On the other hand, speakers of Hindi would not classify [t] and [tʰ] as members of the same class, because they perceive them as different. That is, they are contrastive in Hindi. Similarly, English [tʰ] and [d] are members of different classes because they are contrastive. That is, if you interchange one for the other in a word, you can cause a change in the word's meaning, e.g., *time* [tʰɑɪm] versus *dime* [dɑɪm]).

A class of speech sounds that seem to be variants of the same sound is called a **phoneme.** Each member of a particular phoneme class is called an **allophone,** which corresponds to an actual phonetic segment produced by a speaker. That is, the various ways that a phoneme is pronounced are called allophones.

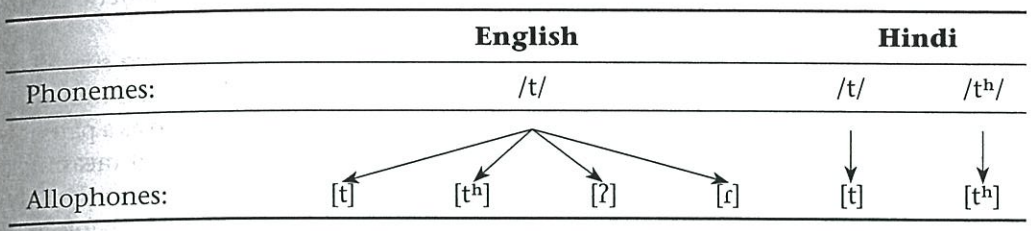
We can say, then, that the 't' sounds in words like *stop, top, little,* and *kitten* all belong to a single class, which we will label by the symbol /t/, characterizing this particular phoneme. By saying that *stop* and *top,* for example, each have the phoneme /t/, we are saying that the sounds [t] and [tʰ] are related—that they are the "same" sound.

[1]The reasons for this may be manifold, including phonetic similarities, phonological patterning, different pronunciations across language varieties, or spelling.

In (5) we see how the phoneme /t/ is related to its allophones in English and how the Hindi phonemes /t/ and /tʰ/ are related to their allophones. In English, [t], [tʰ], [ɾ], and [ʔ] are allophones of the same phoneme, which we can label /t/. In this way, we can say that in English the phoneme /t/ has the allophones [t] as in [stɑp], [tʰ] as in [tʰɑp], [ɾ] as in [lɪɾ̩], and [ʔ] as in [kɪʔn̩].

In Hindi on the other hand, [t] and [tʰ] contrast and are therefore allophones of different phonemes. Note that symbols representing phonemes are written between slashes; this distinguishes them from symbols representing (allo)phones, which are written between square brackets.

(5)



By providing a description like this, linguists attempt to show that the phonological system of a language has two levels. The more concrete level involves the physical reality of phonetic segments, the allophones, whereas phonemes are something more abstract. In fact, linguists sometimes describe phonemes as the form in which we store sounds in our minds. So, phonemes are abstract psychological concepts, and they are not directly observable in a stream of speech; only the allophones of a phoneme are. It is important to note that any sound that is pronounced, then, is an allophone of some phoneme; the phoneme itself is never pronounced.

The phoneme is a unit of linguistic structure that is just as significant to the native speaker as the word or the sentence. Native speakers reveal their knowledge of phonemes in a number of ways. When an English speaker makes a slip of the tongue and says [tʃɑɪn ɹɛk] for *rain check*, reversing [tʃ] and [k], he or she has demonstrated that [tʃ] functions mentally as a single unit of sound, just as [k] does. Recall from File 2.2 that [tʃ] is phonetically complex, consisting of [t] followed immediately by [ʃ]. Yet, since [tʃ] represents the pronunciation of a single phoneme /tʃ/ in English, no native speaker would make an error that would involve splitting up its phonetic components; you will never hear [tɹɑɪn ʃɛk] as a slip of the tongue (see File 9.3).

Knowledge of phonemes is also revealed in alphabetic spelling systems (see File 15.2). For example, English does not have separate letters for [pʰ] and [p]; they are both spelled with the letter *p*. Examples like this show that the English spelling system ignores differences in pronunciation that don't result in meaning distinctions. For the most part, the English spelling system attempts to provide symbols for phonemes, not phonetic segments. In general, alphabetic writing systems tend to be phonemic rather than phonetic, though they achieve this goal with varying degrees of success. As noted in File 2.1, of course, there are multiple ways to represent the same sound (e.g., the [k] sound is written with a <k> in the word *kitten* but with a <c> in the word *cool*). What's crucial here, though, is that both of these spellings represent /k/, and not, for example, the difference between [k] and [kʰ].

3.2.3 Identifying Phonemes and Allophones: The Distribution of Speech Sounds

In order to determine whether sounds in a given language are allophones of a single phoneme or they contrast and are allophones of separate phonemes, we need to consider the

distribution of the sounds involved. The **distribution** of a phone is the set of phonetic environments in which it occurs. For example, nasalized vowels in English occur only in the environment of a nasal consonant. More precisely, a linguist would describe the distribution of English [ĩ], [õ], and so on, by stating that the nasalized vowels always and only occur immediately preceding a nasal consonant. In this book we will mainly be concerned with two types of distribution—contrastive distribution and complementary distribution—though a third distribution, free variation, will also be introduced in the following section.

Let us consider **contrastive distribution** first. Contrastive distribution is simply a case in which the two sounds occur in the same phonetic environment, and using one rather than the other changes the meaning of the word. [p] and [pʰ] in Hindi have a contrastive distribution in Hindi, because they occur in exactly the same phonetic environment but give two different words: [pʰəl] 'fruit' and [pəl] 'moment.'

We can, and did in previous section, determine whether two sounds contrast in a language by taking into account the distribution of sounds in each individual language. We did this by identifying a **minimal pair.** A minimal pair is defined as a pair of words whose pronunciations differ by exactly one sound and that have different meanings. When you find a minimal pair, you know that the two sounds that differ are contrastive. In Hindi, [pʰəl] 'fruit' and [pəl] 'moment' form a minimal pair, showing that [p] and [pʰ] are contrastive in Hindi; [pʰəl] 'fruit' and [bəl] 'strength' also form a minimal pair.

If you try, you can think of many minimal pairs in English, or any other language you know well. For example, the minimal pair [tʰiːm] *team* and [tʰiːn] *teen* shows that [n] and [m] are contrastive in English because they can be used to contrast meaning. But notice that there are no minimal pairs involving [pʰ] and [p] in English; these two sounds are never contrastive with respect to one another. Instead, they are allophones of the same phoneme, /p/.

Consider another example in which two languages make different distinctions using the same set of sounds. In English, [l] and [ɹ] are contrastive, as can be seen from the existence of minimal pairs such as *leaf* [lif] versus *reef* [ɹif], *alive* [əlɑɪv] versus *arrive* [əɹɑɪv], or *feel* [fil] versus *fear* [fiɹ]. In Korean, on the other hand, [l] and [ɾ] are never contrastive.[2] Consider the data in (6).

(6) Korean [l] versus [ɾ] alternations

Citation Form	**Nominative Case**	**Gloss**
[pul]	[puɾi]	'fire'
[mal]	[maɾi]	'language, speech'
[tal]	[taɾi]	'moon'
[kʰal]	[kʰaɾi]	'knife'
[pal]	[paɾi]	'foot'
[sal]	[saɾi]	'flesh'

Notice that in each example, only one English translation is given. For example, the word for 'fire' has two forms: one when it is used by itself, the citation form (such as in the answer to the question "What is the Korean word for 'fire'?"); and one when it is used as

[2]You will notice that the two "r" sounds in English and Korean are not the same phonetically—in English, it is a voiced alveolar retroflex liquid [ɹ], while in Korean, it is a voiced alveolar flap [ɾ] (see Section 2.2.5). Similarly, the exact articulation of /l/ in the two languages is also not identical, though we use the same symbol for both. These phonetic differences are not particularly important here, however, because we are concerned only with the distribution of the two sounds in each language (a phonological question) rather than with the quality of the two sounds (a phonetic question).

the subject of a sentence, the nominative form (such as in "The fire burned brightly"). The final *-i* in the second column of words in (6) indicates that these words are in the nominative case. (For more on these kinds of markers, see Chapter 4.)

The pronunciation of the base word alternates depending on the phonetic context it appears in. An **alternation** is simply a difference between two (or more) phonetic forms that you might otherwise expect to be related. Identifying alternations relies on the assumption that, all else being equal, the same meaning should be expressed by the same sounds—when we find different pronunciations of the same word that are systematically linked to particular grammatical contexts, we have an alternation. In Korean, the base word *fire* is expressed alternately by the sounds [pul] and [puɾ].

These words have alternating pronunciations depending on what other elements they appear with. How do you know which form of the word to use? Consider the distribution of the sounds [l] and [ɾ] in Korean. In the words listed in (6), we see that [l] and [ɾ] do not occur in the same phonetic environment. Specifically, as you can see from the data in (6), you use an [l] when the sound is the last sound in the word (is in "word-final" position), but you use [ɾ] when the sound is between two vowels (is in "intervocalic" position). In fact, if you were to look at all of Korean, you would find that [ɾ] appears only between two vowels, while [l] never appears in that position. Meanwhile, [l] can appear at the ends of words, but [ɾ] never does.[3] These two observations mean that if someone gave you the frame [tɑ_] in Korean, you could tell them whether an [ɾ] or an [l] goes in the blank. Because where these two sounds occur is predictable in Korean, only an [l] can go in the blank. Notice that you cannot do the same thing in English: either an [l] or an [ɹ] could go in the blank to form a possible English word: *tall* [tɑl] or *tar* [tɑɹ].

The difference that has been illustrated here between English and Korean is that in English, [l] and [ɹ] are in contrastive distribution, while in Korean, [l] and [ɾ] are in **complementary distribution.** Sounds in complementary distribution are considered to be allophones of the same phoneme. To understand better what we mean by complementary distribution, think about what the term complementary means: two complementary parts of something make up a whole. For example, the set of people in your class at any given moment can be divided into the set of people who are under 5′5″ tall and the set of people who are 5′5″ tall or taller. These two sets of people complement each other. They are mutually exclusive (one person can't simultaneously be both shorter and taller than 5′5″), but together they make up the whole class.

The Korean sounds [l] and [ɾ] are in complementary distribution because they appear in different sets of environments: [ɾ] occurs between vowels, and [l] occurs word-finally. Given our assumptions about alternations—that the same word should be expressed by the same sounds—we can hypothesize that even though [l] and [ɾ] are phonetically different phones in Korean, they are allophones of a single phoneme. We can represent that single phoneme as /l/. At this point, it probably seems like an arbitrary choice as to why /l/ should be the phoneme and not /ɾ/; we will talk more about how to make this choice in File 3.5.

If two sounds are in complementary distribution in a language, there will never be a minimal pair that uses them to distinguish two words; that is, they are not contrastive. Furthermore, the appearance of one allophone or the other will always be predictable, as we saw above with the frame [tɑ_]. You can predict that [ɾ] but not [l] will appear between vowels in any word in Korean, and that [l] but not [ɾ] will appear word-finally—even if you have never studied Korean. Similarly, in the Kikamba example in 3.2.1, [k] and [g] are in complementary distribution: [g] but not [k] occurs after [ŋ], and [k] but not [g] occurs everywhere else. This kind of prediction is a powerful tool in helping phonologists understand the structure of languages.

[3] In some modern words that have come into Korean from other languages, [l] can also appear at the beginning of words. But it never appears as the only segment between two vowels.

What's particularly interesting about this (and all other) phonological distributions is that it represents actual knowledge that native speakers have. For example, if you give a native speaker of Korean the new (nonsense) word *moladam* and ask them to say it out loud, they will say it with an [ɾ] between the two vowels, and not an [l]. Of course, this is not something that anyone has explicitly taught them (especially since they have never seen this word before), but the distribution of sounds is one of the things that you know when you know a language (see File 1.2).

Consider another linguistic example, namely, the distribution of the English sounds [p] and [pʰ] shown in (7).

(7)

spat	[spæt]	*pat*	[pʰæt]
spool	[spul]	*pool*	[pʰul]
speak	[spik]	*peek*	[pʰik]

As you can see in the English words in (7), [p] and [pʰ] do not occur in the same phonetic environment. As a result, there are no minimal pairs involving a [p]-[pʰ] contrast. In fact, the phones are in complementary distribution: [p] occurs after [s] but never word-initially, while [pʰ] occurs word-initially but never after [s]. Since these sounds appear in different phonetic environments, there can be no pair of words composed of identical strings of sounds except that one has [p] and the other has [pʰ]. We know that phones that are in complementary distribution are allophones of a single phoneme. In this case, [p] and [pʰ] are both allophones of the phoneme we can represent as /p/. Furthermore, the appearance of [p] or [pʰ] in a given context is predictable. For example, we can predict that the allophone [pʰ] (but never [p]) will appear in word-initial position. So even in words not listed in (7) such as *pot* or *pin*, we know that [pʰ] rather than [p] will occur at the beginning of a word.[4] Similarly, we can predict that [p] (but never [pʰ]) will follow [s] in other words such as *spot* and *spin*.

We can summarize the difference between sounds that are contrastive (e.g [p] and [pʰ] in Hindi or [l] and [ɹ] in English) and sounds that are allophones of the same phoneme (e.g. [p] and [pʰ] in English or [l] and [ɾ] in Korean) as shown in (8).

(8)

	Contrastive	**Allophonic**
Relation to phonemes	Allophones of separate phonemes	Allophones of the same phoneme
Predictability of distribution	Unpredictably distributed	Predictably distributed
How you can tell	Contrastive distribution; minimal pairs	Complementary distribution

3.2.4 Free Variation

Most phonological distributions can be described as either contrastive or complementary. Remember that the hallmark of a contrastive distribution is that you can't predict which of two (or more) sounds belongs in a certain context, because each will produce a different but meaningful word, while the hallmark of a complementary distribution is that you can predict which of two sounds belongs in any given context.

[4] In point of fact, this is true not just at the beginning of a word but at the beginning of any stressed syllable. That is, in English, [pʰ] but not [p] can appear as the first consonant of a stressed syllable.

In some contexts, however, more than one pronunciation of a given sound may be possible without changing the meaning of the word. In these cases, you may not be able to predict exactly which sound will occur, but the choice does **not** affect the meaning of the word. Consider, for example, the pronunciations of some English words in (9) (note that [p̚] represents an unreleased voiceless bilabial stop).

(9)

leap	[lip]	*leap*	[lip̚]
soap	[soʊp]	*soap*	[soʊp̚]
troop	[trup]	*troop*	[trup̚]
happy	[hæpi]	—	*[hæp̚i]

These words show that [p] and [p̚] both share some of the same phonetic environments; specifically, they can both appear at the ends of words. Unlike the case of English [b] versus [pʰ], or [m] versus [n], however, there are no minimal pairs involving these sounds in the language. Why not? Although there are pairs of words in (9) that differ in only one sound, none of these words contrast in meaning. Thus, the choice between [p] and [p̚] in *leap, soap,* and *troop* does not make a difference in meaning; that is, the sounds are noncontrastive. Rather, they are interchangeable in word-final position. Sounds with this type of patterning are considered to be in **free variation.** To a native speaker, sounds like [p] and [p̚] that are in free variation are perceived as being the "same" sound. We can conclude that they are allophones of the same phoneme, because they are perceived as the same and do not serve to distinguish the meanings of words.

Because [p] and [p̚] can occur in the same environment, they are in what is called **overlapping distribution;** they can occur in the same environment. Sounds that are in contrastive distribution and sounds that are in free variation are therefore both considered to have an overlapping distribution; only sounds that are in complementary distribution do not overlap. For example, in English, the sounds [d] and [t] are in overlapping distribution because they can occur in the same phonetic environment. The words *lid* and *lit* form a minimal pair, and both [d] and [t] can occur after [lɪ]; that is, the environment [lɪ_] is one where [d] and [t] overlap. Similarly, [t] and [t̚] have an overlapping distribution because they can also both occur after [lɪ_], as two different pronunciations of the word *lit.* The difference between [d] and [t] on the one hand, and [t] and [t̚] on the other, is that interchanging [d] and [t] changes the meaning of the words, while interchanging [t] and [t̚] does not.

We can thus extend our table in (8) to include the characteristics of free variation, as shown in (10).

(10)

	Contrastive	**Allophonic**	**Free Variation**
Relation to phonemes	Allophones of separate phonemes	Allophones of the same phoneme	Allophones of the same phoneme
Predictability of distribution	Unpredictable	Predictable	Unpredictable
How you can tell	Contrastive distribution; minimal pairs	Complementary distribution	Overlapping distribution with no difference in meaning

FILE 3.3

Phonological Rules

3.3.1 Phonological Rules

In File 3.2, we discussed the fact that phonemes and (allo)phones belong to different levels of structure in language—that is, phonemes are abstract mental entities, and phones are physical events. In this file we consider the connection between these two levels.

The mapping between phonemic and phonetic elements is accomplished using **phonological rules** (recall from Section 1.2.3 that a rule of grammar expresses a pattern in a language). A speaker's knowledge of phonological rules allows him or her to "translate" phonemes into actual speech sounds; knowledge of these rules forms part of the speaker's linguistic competence. This change from the phonemic form, also called the **underlying form,** to the actual phonetic form of a word by means of phonological rules can be represented with the diagram in (1).

(1) phonemic form
⇓
rules
⇓
phonetic form

As an example, consider the English words *seat* /sit/ and *loot* /lut/. These words have a final /t/ sound in their phonemic form which is often pronounced as [t], at least in careful speech.[1] Now compare the pronunciation of /t/ in the forms of these words when the suffix *-ed* is added, as shown in (2).[2]

(2)	*seat*	[sit]	*seated*	[sitɪd]
	loot	[lut]	*looted*	[lutɪd]

As these transcriptions show, the phoneme /t/ is pronounced as the flap [ɾ] when it occurs between two vowels, specifically when the preceding vowel is stressed and following one is unstressed. We can state this observation about English in terms of a descriptive rule, as in (3). (We will be modifying this rule later in this file since, as we know from our discussion of /t/ in File 3.2, there are more allophones of /t/ than just [t] and [ɾ].)

---

[1]In linguistic analysis, we often have to distinguish between "careful" and "casual" speech. Careful speech is when a speaker speaks more slowly and clearly than usual, while casual speech is when a speaker speaks more quickly and with more co-articulation (see Section 2.2.6) than usual. Of course, these are really endpoints on a scale of speech styles (see File 10.1), and people actually talk at many different styles in between. The pronunciations listed in (2) are used in all but the most careful styles of speech.

[2]Here and throughout this file, we use a fairly broad transcription system, recording phonetic detail only if relevant for the segments under discussion.

(3) /t/ is pronounced as [ɾ] after a stressed vowel and before an unstressed vowel
[t] everywhere else

Notice that a phonological rule has three parts: the sound(s) affected by the rule, the environment where the rule applies, and the result of the rule. In the rule in (3), /t/ is affected by the rule. The rule applies when /t/ occurs after a stressed vowel and before an unstressed vowel. The result of the rule is that /t/ is "flapped"; that is, it is pronounced as [ɾ].

We can write this rule using shorthand of the form X→Y / C _ D. Here, 'X' is the sound that is affected by the rule, 'Y' is the result of the application of the rule, and 'C _ D' is the environment in which the rule applies. 'C _ D' is also called the **conditioning environment.** By "C _ D," we mean that C comes before the sound affected by the rule and D comes after it; the blank represents where the sound that is affected by the rule appears. You can read these rules in the following way: "X becomes Y when it comes after C and before D." Thus, if you see CXD, you know that it will become CYD when the rule applies. So, for the rule in (3), we would write:

(4)

X	→	Y	/	C	_	D
/t/	→	[ɾ]	/	stressed vowel	_	unstressed vowel
/t/	→	[t]	/	everywhere else		

In (5), we illustrate how the phonetic forms of the examples in (2) are derived from the phonemic forms. You can see that the rule applies to the phonemic forms of *seated* and *looted* because /t/ occurs in the correct context for the rule, as defined in (4). However, in *seat* and *loot* the rule does not apply since /t/ does not occur in the context for flapping.

(5)

phonemic form:	/sit/	/sit + ɪd/	/lut/	/lut + ɪd/
apply rule	—	siɾɪd	—	luɾɪd
phonetic form	[sit]	[siɾɪd]	[lut]	[luɾɪd]

The derivation in (5) illustrates what happens in speaking. In listening, a hearer reverses this process: he or she perceives the phonetic form of an utterance, then sends it "backwards" through the phonological rules, and finally obtains a phonemic form that matches a form stored in memory.

3.3.2 Natural Classes

In the previous section, we observed that the phoneme /t/ can be pronounced as a flap. In fact, it is not only /t/ that can be flapped, but /d/ as well, as shown in (6). For speakers of American English, the middle /d/ in *seeded* and *seated* are both pronounced as a flap, making these words sound very similar, if not identical.

(6) *seed* [sid] *seeded* [siɾɪd]

Since /d/ also undergoes flapping, we can modify our rule in (4) as follows:

(7) /t, d/ → [ɾ] / stressed vowel __ unstressed vowel

Given the observation that both /t/ and /d/ are subject to the same phonological rule, we can now ask whether it is random chance that these two sounds are both subject to the flapping rule or if there is a more principled reason for their similar patterning.

To answer this question, let's consider the articulatory descriptions of the two sounds.

(8) /t/ voiceless alveolar (oral) stop
/d/ voiced alveolar (oral) stop

Not only are both sounds alveolar stops, they are the **only** oral alveolar stops in English.[3] Therefore, we can make the description more general by removing some of the properties.

(9) /t, d/ alveolar (oral) stop

With respect to English, saying "alveolar (oral) stop" is the same as saying /t/ and /d/. These two sounds are the only oral phonemes in English that are produced by stopping the flow of air at the alveolar ridge. Thus, they are the **natural class** of alveolar (oral) stops. A natural class is a group of sounds in a language that share one or more articulatory or auditory property, **to the exclusion of all other sounds in that language.** That is, in order for a group of sounds to be a natural class, it must include all of the sounds that share a particular property or set of properties, and not include any sounds that don't.

All of the properties used in Files 2.2, 2.3, and 2.4 to describe individual sounds can also be used to describe natural classes. For example, in the English vowels the monophthongs [i, u] and the first part of the diphthongs [eɪ] and [oʊ] are tense vowels, and there are no other tense vowels in English. Thus, these four vowels are members of the natural class of tense vowels in English. Likewise, the consonants [k, g, ŋ] are all described as velar consonants, and they are the only velar consonants used in English; thus they constitute the natural class of velar consonants in English. Natural classes can be used to describe both the sounds affected by a rule **and** the environments where a rule applies.

In talking about groups of sounds, we must use a few properties in addition to those needed to describe individual sounds. One new property that we will need is **sibilant.** Sibilants are segments that have a high-pitched, hissing sound quality. The natural class of sibilants in English is [s, ʃ, tʃ, z, ʒ, dʒ].

In addition, if you look at the consonant chart on the last page of this book, you will notice that the only labiodental consonants in English are the fricatives [f] and [v], while the bilabial fricative slots are left empty. In many situations it is advantageous to refer to [f] and [v] together with [p, b, m, w] and [ʍ] as belonging to the same natural class. For this purpose we use the property **labial.**

Another property used to describe natural classes divides the segments into two groups, **obstruents** and **sonorants.** Obstruents are produced with an obstruction of the airflow. The sounds in this category are stops, fricatives, and affricates. Sonorants, on the other hand, are segments produced with a relatively open passage for the airflow. Sonorant segments include nasals, liquids, glides, and vowels. Thus, the class of labial obstruents in English is [p, f, b, v], while the class of labial sonorant consonants is [m, w, ʍ].[4] The class of labial consonants is the union of both sets: [p, f, b, v, m, w, ʍ]. As we will see, being able to divide consonants into obstruents and sonorants is quite useful in stating phonological rules.

---

[3]We say that /t/ and /d/ are the only **oral** alveolar stops because nasal /n/ can also be classified as an alveolar nasal stop.

[4]As already mentioned, the class of sonorants also includes vowels, because they do not have an obstruction of airflow. Some linguists treat rounded vowels as "labial," in which case the entire class of English labial sonorants would be [m, w, ʍ, u, ʊ, o, ɔ], and the entire class of English labials would be [p, f, b, v, m, w, ʍ, u, ʊ, o, ɔ].

3.3.3 Types of Phonological Rules

Every language has many phonological rules. In addition to seeing that phonological rules apply to natural classes of segments, we can classify phonological rules according to the kind of process that they involve. Seven major types of processes are discussed here, along with examples from the phonology of English and other languages.

a. Assimilation. Rules of **assimilation** cause a sound (or gesture) to become more like a neighboring sound (or gesture) with respect to some phonetic property. In other words, the segment affected by the rule assimilates or takes on a property from a nearby (often adjacent) segment. Rules of assimilation are very common in languages. An example of assimilation is the pronunciation of the prefix *un-* in English. Words like *unbelievable, unstable,* and *unclear* are often pronounced [ʌmbəlivəbl̩], [ʌnsteɪbl̩], and [ʌŋkliɹ]. That is, the nasal /n/ is often pronounced as a bilabial nasal when it occurs before a bilabial sound, as in *unbelievable,* and as a velar nasal when it occurs before a velar sound, as in *unclear.* This is called Nasal Place Assimilation because the nasal /n/ changes its place of articulation:

(10) **Nasal Place Assimilation** (English): An alveolar nasal assimilates to the place of articulation of a following consonant.

Thus, when a sound having the properties alveolar and nasal immediately precedes a labial consonant, this rule causes the alveolar nasal to take on the property labial (thereby replacing its specification for alveolar).

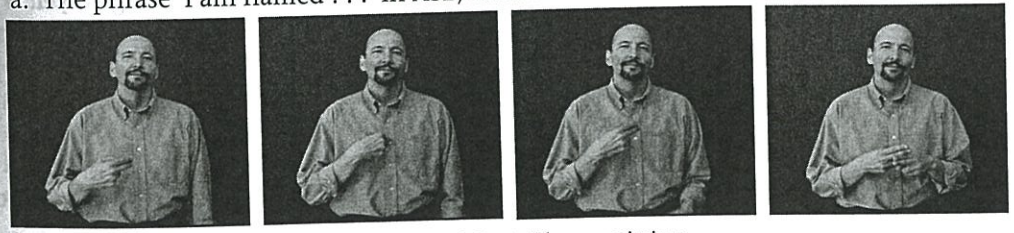
We can see a similar sort of phenomenon taking place across word boundaries in certain ASL handshapes. We will consider the handshape that is used in the sign ME, which is a pointing index finger, as shown in (11).

(11) The unassimilated sign for ME in ASL

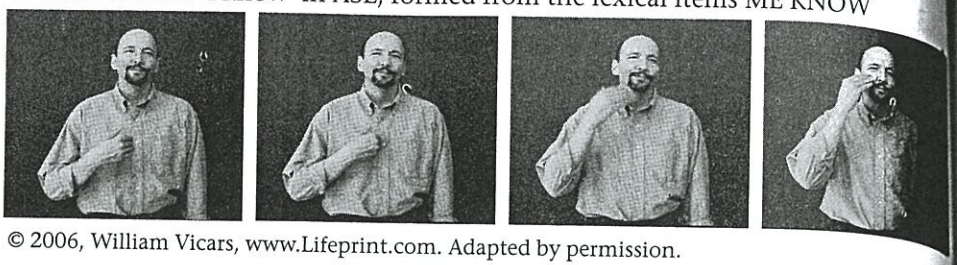
The sign ME may take on features of other handshapes, however, depending on the sign that follows it. For example, in order to say "I am named . . . ," a speaker of ASL would sign "ME NAME . . . ." In order to say "I know," a speaker of ASL would sign "ME KNOW." The signs NAME and KNOW have different handshapes: NAME is articulated with two fingers (index finger and middle finger) extended; KNOW is articulated with a bent hand and all four fingers extended.

When the sign ME is produced before one of these other words, it can take on the handshape of the word that follows it, as shown in (12).

(12) a. The phrase 'I am named . . .' in ASL, formed from the lexical items ME NAME



b. The sentence 'I know' in ASL, formed from the lexical items ME KNOW



Notice that in both (12a) and (12b), the signer touches his chest with his hand facing the same way as he would in the unassimilated form. That is, the place of articulation, the orientation, the movement, and the non-manual marker for ME do not change. But the handshape used for ME in (12a) is the handshape of NAME, and the handshape used in (12b) is the handshape of KNOW.

Another assimilation process is **palatalization.** Palatalization refers to a special type of assimilation in which a consonant becomes like a neighboring palatal. For example, when American English speakers say *Did you?* rapidly, they very often pronounce it as [dɪdʒu]. The sounds [d] (the alveolar stop from the end of *did*) and [j] (the palatal glide from the beginning of *you*) combine to form the palatal affricate [dʒ]. In this case, the palatal nature of the glide has been assimilated by the stop, making it a palatal affricate. Front vowels such as [i] and [e] also cause this change. The most common types of palatalization occur when alveolar, dental, and velar stops or fricatives appear before a front vowel. So the following are all common types of palatalization: [t] → [tʃ]; [d] → [dʒ]; [s] → [ʃ]; [k] → [tʃ]; [g] → [dʒ]. While there are variants on palatalization, and other sounds can be palatalized, the main things to look for are a sound becoming a palatal and/or a phonological rule conditioned by a front vowel.

The rules of assimilation that we've discussed so far cause sounds to assimilate to **adjacent** sounds. This is a common way that assimilation occurs. However, long-distance assimilation also exists, and a relatively common type of long-distance assimilation is called **vowel harmony.** This typically causes all the vowels in a word to "harmonize" or agree in some property such as rounding or backness.

Finnish has a common type of vowel harmony rule, which can be stated as follows:

(13) **Vowel harmony** (Finnish): A back vowel becomes front when preceded by a front vowel in the same word.

By this rule, Finnish words have, with few exceptions, either all front vowels or all back vowels, but not both in the same word. We can see the vowel harmony rule in action when a suffix is added to the end of a word. In this case, the suffix vowel changes to match the quality of vowels in the word. For example, the suffix meaning 'in' has the form [-ssɑ] when added to a word where the last vowel is back, as in [tɑlo] 'house,' [tɑlossɑ] 'in the house.' However, the suffix takes the form [-ssæ] when it attaches to a word with a final front vowel, as in [metsæ] 'forest,' [metsæssæ] 'in the forest.' In cases like this, we can say that the vowel of the suffix harmonizes with, or assimilates to, the preceding vowel.

b. Dissimilation. Unlike assimilation, which makes sounds more similar, rules of **dissimilation** cause two close or adjacent sounds to become less similar with respect to some property, by means of a change in one or both sounds. An example of dissimilation in Greek is the following:

(14) **Manner dissimilation** (Greek): A stop becomes a fricative when followed by another stop.

For example, in fast speech especially, the form /epta/ 'seven' can be pronounced as [efta], and /ktizma/ 'building' can be pronounced as [xtizma] ([x] is a voiceless velar fricative).

c. Insertion. Phonological rules of **insertion** cause a segment not present at the phonemic level to be added to the phonetic form of a word. An example of this kind of rule from English is voiceless stop insertion:

(15) **Voiceless stop insertion** (English): Between a nasal consonant and a voiceless fricative, a voiceless stop with the same place of articulation as the nasal is inserted.

Thus, for instance, the voiceless stop insertion rule may apply to the words *dance* /dæns/ → [dænts], *strength* /stɹɛŋθ/ → [stɹɛŋkθ], and *hamster* /hæmstɻ/ → [hæmpstɻ].

d. Deletion. Deletion rules eliminate a sound that was present at the phonemic level. Such rules apply more frequently to unstressed syllables and in casual speech. English examples include:

(16) **/h/-Deletion** (English): /h/ may be deleted in unstressed syllables.

The /h/-deletion rule would apply to a sentence such as *He handed her his hat* /hi hændəd hɻ hɪz hæt/ to yield [hi hændəd ɻ ɪz hæt]. Deletion is common in fast speech because it saves time and articulatory effort. Sounds like [h] that are not very perceptible are often the "victims" of deletion because speakers can save time and effort by deleting them without sacrificing much information. That is, the listener may not be relying on these sounds in order to understand what the speaker is saying.

e. Metathesis. Rules of **metathesis** change the order of sounds. In many instances, sounds metathesize in order to make words easier to pronounce or easier to understand. In Leti, an Austronesian language, consonants and vowels switch places when a word that ends in a consonant is combined with a word that starts with two consonants. The last two sounds in the first word trade places to avoid having three consonants in a row.

(17) **CV metathesis** (Leti): When three consecutive consonants occur, the first consonant trades places with the preceding vowel.

By this rule, /danat + kviali/ 'millipede' undergoes metathesis to become [dantakviali], and /ukar + ppalu/ 'index finger' becomes [ukrappalu]. On the other hand, /ukar + lavan/ 'thumb' does not undergo metathesis and so is pronounced as [ukɑrlɑvɑn] because there are not three consecutive consonants.

f. Strengthening. Rules of **strengthening** (also called fortition) make sounds stronger. The rule of English aspiration, as stated below, provides an example:

(18) **Aspiration** (English): Voiceless stops become aspirated when they occur at the beginning of a stressed syllable.

The pronunciations of *pat* /pʰæt/ and *top* as [tʰɑp], as discussed in File 3.2, illustrate the application of the English aspiration rule. Aspirated stops are considered to be stronger sounds than unaspirated stops because the duration of voicelessness is much longer in aspirated stops (since it extends through the period of aspiration).

g. Weakening. Rules of **weakening** (also called lenition) cause sounds to become weaker. The "flapping" rule of English, discussed in 3.3.1, is an example of weakening. [ɾ] is considered to be a weaker sound than [t] or [d] because it is shorter and it obstructs air less.

(19) **Flapping** (English): An alveolar (oral) stop is realized as [ɾ] when it occurs after a stressed vowel and before an unstressed vowel.

Note that voicing assimilation is involved in the change of /t/ to [ɾ]: the /t/ takes on the "voicedness" of the vowels surrounding it.

3.3.4 Multiple Rule Application

To this point we have seen examples where only one phonological rule applies. In reality there is often more than one change that occurs between a given phonemic form and the phonetic output. To illustrate this let's look at how plural nouns are formed in English. When you learned to write in English, you learned that the way to make most nouns plural is to add an <s>, which is usually pronounced [z]. There are actually three different phonetic forms of the English plural marker: [s], [z], and [əz], as in the words *cats* [kæts], *dogs* [dɔgz], and *foxes* [fɑksəz]. We need only one phonemic form for the plural marker if we use two rules to derive the phonetic forms.

Let's assume that the single phonemic form of the plural marker is /-z/, rather than /-s/ or /-əz/.[5] Why do words like *fox, ditch, bush, orange, rouge,* and *maze* have [-əz] as their plural suffix instead of just keeping the phonemic form /-z/? The answer lies in the last sound of each of these words—they are all sibilant consonants. Notice that the plural marker is also a sibilant /-z/. Because of the high-pitched hissing sound, it is very difficult to hear two sibilants that are next to each other. Try saying [fɑksz], [dɪtʃz], [bɹɪdʒz], etc., and you will get the idea. This difficulty is remedied by inserting a schwa between the two sibilants.

(20) **Schwa insertion** (English): Insert [ə] between two sibilants.

The schwa insertion rule makes these plurals more distinct by separating the two sibilants: [fɑksəz], [dɪtʃəz], [bɹɪdʒəz], etc. But why do we still need two different plurals, [-s] and [-z], for words that do not end in sibilants?

Try to pronounce [kætz] or [dɔgs], in which the voicing qualities of the final two consonants differ. You will probably find that it is difficult to produce a consonant cluster if one consonant is voiced and the other is voiceless. This may be one reason that the plural marker changes its voicing specification to match the sound it follows.

(21) **Voicing assimilation** (English): /-z/ takes on the voicing specification of the preceding sound.

The voicing assimilation rule takes a phonemic form like /kæt-z/ and turns it into the pronounceable form [kæts]. With these two rules acting together, we can derive the plural for any English noun (except, of course, for "special" or irregular plurals like *oxen, cacti,* or *cherubim*). Starting with the phonemic form, we apply each rule in turn. But how do we know which rule to apply first? In most cases, phonological rules must apply in a particular order in order to arrive at the correct phonetic form. The best way to decide on the ordering of the rules is by trial and error: choose an order and try to derive the correct phonetic forms. In (22), we can see how these **derivations** work: (22a) gives one possible ordering, with schwa insertion before the voicing assimilation, while (22b) gives a derivation with the other order. Which derivation gives the correct phonetic form of the plurals?

(22) Sample derivations of English plurals

a. Schwa insertion ordered before voicing assimilation

phonemic form:	/kæt-z/	/dɔg-z/	/fɑks-z/	/bɹɪdʒ-z/
schwa insertion:	—	—	fɑksəz	bɹɪdʒəz
voicing assimilation:	kæts	—	—	—
phonetic form:	[kæts]	[dɔgz]	[fɑksəz]	[bɹɪdʒəz]

[5]This assumption is in fact correct for English; the arguments for picking this form, however, are beyond the scope of this discussion.

b. Voicing assimilation ordered before schwa insertion

phonemic form:	/kæt-z/	/dɔg-z/	/fɑks-z/	/bɹɪdʒ-z/
voicing assimilation:	kæts	—	fɑkss	—
schwa insertion:	—	—	fɑksəs	bɹɪdʒəz
phonetic form:	[kæts]	[dɔgz]	*[fɑksəs]	[bɹɪdʒəz]

As you can see in (22b), ordering the voicing assimilation before the schwa insertion provides an incorrect phonetic form for *foxes* *[fɑksəs] instead of the correct [fɑksəz]. Therefore, we know that the schwa insertion rule has to apply before the voicing assimilation rule in order to derive the correct phonetic forms.

3.3.5 Obligatory and Optional Rules

Notice that phonological rules may be **obligatory** or **optional.** Obligatory English rules include aspiration, vowel nasalization, vowel lengthening, and liquid and glide devoicing. Such a rule always applies in the speech of all speakers of a language or dialect having the rule, regardless of style or rate of speaking. The effects of obligatory rules are often very subtle and difficult to notice, but they are an important part of a native accent. For instance, it may be difficult to tell that a vowel is nasalized in English, but not applying the rule of vowel nasalization would make someone sound like a non-native speaker of English.

The existence of obligatory rules is what causes people to have foreign accents. It is easier to learn the rules of a new language than to "turn off" the obligatory rules of your native language. The very fact that we are often unaware of these rules causes us to apply them when they are not appropriate. When speakers of American English learn other languages, they often apply rules such as flapping and vowel reduction, in addition to the phonotactic constraints discussed in File 3.1, even though the other language may not have these rules.

Optional phonological rules, on the other hand, may or may not apply in any given utterance. Optional rules are responsible for variation in speech; for example, we can pronounce *He handed her his hat* as [hi hændəd hɨ hɪz hæt] or [hi hændəd ɨ ɪz hæt], depending on whether the /h/-deletion rule is applied or not. The use of optional rules depends in part on rate and style of speech.