3 More on Phonemes

3.1 Phonemic Analysis and Writing

The question of phonemicization is in principle independent from the question of writing; that is, there is no necessary connection between letters and phonemes. For example, the English phoneme /eɪ/ can be spelled in quite a few ways: ***say*** /seɪ/, ***Abe*** /eɪb/, ***main*** /meɪn/, ***beige*** /beɪʒ/, ***reggae*** /ˈrɛgeɪ/, *H* /eɪtʃ/. Indeed, there are languages (for example, Mandarin Chinese) that are written with symbols that do not correspond to phonemes at all.

Obviously, there is at least a loose connection between alphabetic letters and phonemes: the designers of an alphabet tend to match up the written symbols with the phonemes of a language. Moreover, the conscious intuitions of speakers about sounds tend to be heavily influenced by their knowledge of spelling – after all, most literate speakers receive extensive training in how to spell during childhood, but no training at all in phonology.

Writing is prestigious, and our spoken pronunciations are sometimes felt to be imperfect realizations of what is written. This is reflected in the common occurrence of **spelling pronunciations**, which are pronunciations that have no historical basis, but which arise as attempts to mimic the spelling, as in *often* [ˈɔf͡tən] or *palm* [pɑl̪m].

In contrast, most linguists feel that spoken language is primary, and that written language is a derived system, which is mostly parasitic off the spoken language and is often rather artificial in character. Some reasons that support this view are that spoken language is far older than writing, it is acquired first and with greater ease by children, and it is the common property of our species, rather than of just an educated subset of it.

One of the practical applications of phonology (especially, phonemic analysis) is the design of alphabets for languages heretofore unwritten. It is widely agreed that an alphabet will be most useful and easy to use if the letters and sounds are in one-to-one correspondence. What is not commonly realized is that at the level of phonetic detail – with all the allophones considered – every language contains thousands of sounds. What is really needed in a good alphabet is a system in which it is letters and *phonemes* that correspond one-to-one; this will permit all meaning-bearing distinctions to be reflected in the spelling, and it avoids the

needless effort (and perhaps confusion) that arises in spelling allophonic variation. Advocates of "phonetic spelling" are thus really advocates of "phonemic spelling." Most of the many alphabets designed for hitherto unwritten languages in the last few decades have been phonemic.

3.2 The Psychological Reality of the Phoneme

Chapter 2 covered a fairly mechanical procedure for extracting a phonemic analysis from the data of a language. However, the goal of phonology is not really to contrive useful procedures for arranging linguistic data; rather, we wish to arrive at a scientific account of the knowledge of language internalized by speakers. We can ask: is it legitimate to suppose that speakers actually produce and perceive language (at an unconscious level) in terms of phonemes? A fair amount of experimentation has been conducted on this issue, and the practical experience of linguists over many years likewise seems relevant.

We will discuss this question in three parts. Recall that the most fundamental phenomenon concerns phonetic differences that are contrastive vs. noncontrastive. Evidence that this distinction is part of speakers' knowledge is discussed in §3.2.1 Beyond this, there is the idea that the principle of contrast is expressed in the grammar by grouping together the non-contrasting sounds into phonemes; this is covered in 2.2. Lastly, there is the system of phonological rules which derives the allophones in their proper contexts; this is discussed in 2.3.

3.2.1 Audibility of fine distinctions

There is little doubt that contrastiveness plays a major role in the perceptions of language users. The auditory processing apparatus is "tuned" through experience to be able to extract precisely those phonetic distinctions that are phonemic in the perceiver's own language.

The practical experience of linguists and other language users attests to this. Suppose we are dealing with two sounds that are phonetically rather close. Suppose further that the two sounds are heard by two different listeners. For listener A, the two sounds are contrastive, serving to distinguish words in her language. The two sounds also occur in listener B's language, but are allophones, and are not contrastive. What usually happens in such a situation is this: A can hear the difference between the two sounds with perfect ease, but B has great difficulty.

On various occasions I have played the role of both A and B. Here are some examples.

On one occasion when I was listener B, A was a speaker of a dialect of Bengali in which dental stops (tongue tip touches upper teeth) contrast with alveolar

stops (tongue tip touches alveolar ridge), as in the following minimal and near-minimal pairs:

[t̪an]	'(vocal) tune'	[tan]	'pull!'
[sat̪]	'seven'	[sat]	'sixty'
[d̪an]	'donation'	[dan]	'right (hand)'
[d̪in]	'day'	[dim]	'egg'

Attempting to transcribe A's speech, I found that, despite extensive practice, I was unable to learn to hear the Bengali dental/alveolar distinction. This amused A, who as a native speaker found the distinction to be utterly obvious. (In desperation I tried guessing alveolar on every occasion, a strategy which does better than chance, but my colleague quickly detected this.)

An important additional fact here is that my own native language, a variety of American English, does include both dental and alveolar consonants. The dentals occur as allophones of the alveolars, and are derived by a rule that replaces alveolars by dentals before dental fricatives (this is a generalized version of the rules seen for /n/ and /l/ dentalization in chapter 2).

/eɪtθ/	[eɪt̪θ]	*eighth*	/eɪt ðə/	[eɪt̪ ðə]	*ate the . . .*
/wʊd θɪŋk/	[wʊd̪ θɪŋk]	*would think*	/sɛd ðə/	[sɛd̪ ðɪs]	*said this*
/tɛnθ/	[tɛn̪θ]	*tenth*	/ɪn ðə/	[ɪn̪ ðə]	*in the . . .*

The point is that my inability to hear the dental/alveolar distinction is not due to a lack of experience with dentals. Rather, it is because my native language does not have a phonemic contrast between alveolars and dentals.

I have also played the role of A, as a native speaker of a dialect of English that contrasts /ɑ/ with the lower mid back rounded vowel /ɔ/. In this dialect, there are minimal pairs such as the following:

caught	['kɔt]	*cot*	['kɑt]
Kaun	['kɔn]	*con*	['kɑn]
paw	['pɔ]	*Pa*	['pɑ]
auto	['ɔɾoʊ]	*Otto*	['ɑɾoʊ]

This dialect is spoken by perhaps half of the American population, and coexists with a historically innovating dialect in which all occurrences of the old phoneme /ɔ/ have been replaced by /ɑ/.[1] Thus in the new dialect, *caught* is [kɑt], *Kaun* is [kɑn], *paw* is [pɑ], and *auto* is ['ɑɾoʊ], all homophonous with their counterparts that had [ɑ] from earlier on.

Speakers of this newer dialect often claim that hearing the [ɑ]/[ɔ] distinction is quite difficult for them. In contrast, speakers of the older dialect, which makes a distinction, feel that it would be very strange *not* to be able to hear it.

[1] A map showing where the two dialects are spoken may be examined at www.ling.upenn.edu/phono_atlas/maps/Map1.html (Telsur Project, University of Pennsylvania).

Such stories are easily multiplied. They have also been backed up by extensive experimental work, which can now actually trace the time course whereby humans change from unbiased perceivers to phonological perceivers, who fluently detect the contrasts of their own language, but not that of others. Strikingly, much of the process appears to take place in infancy: ten-month-olds already show a strong tendency to hear only those phonetic distinctions that are present in the ambient language.

To summarize: the contrastiveness of two phonetically similar sounds leads speakers of a language or dialect that has the contrast to focus their perceptual attention on the contrasting sounds, and fail to hear other distinctions. In a phoneme-based approach to phonology, we would interpret this as indicating that native speakers hear the differences between phonemes, but not between allophones.

There is a practical consequence of all this: if you are trying to elicit new data from a native speaker, and are not a native speaker yourself, it is likely that you will have trouble in hearing the finer phonemic distinctions that do not occur in your own language. In such cases, you can often improve the quality of your data by having your consultant *help you listen*, using the method of **keywords**. If, for instance, you speak a dialect of English that has no contrast between /ɑ/ and /ɔ/, and you were eliciting from me a series of words that contained either /ɑ/ or /ɔ/, I could help you out in identifying the correct vowels – my phonemic distinction (and lifelong experience in hearing it) makes it highly detectable for me. This can be achieved by setting up keywords which unambiguously identify the target vowel. For example, for English /ɑ/ vs. /ɔ/, one could establish *shah* [ʃɑ] as the keyword for /ɑ/ and *saw* [sɔ] as the keyword for /ɔ/. Then, whenever the identity of a vowel is not clear, it can be determined by asking the consultant "vowel of *shah* or vowel of *saw*?"

3.2.2 The notion of "same sound"

The evidence above focuses on the most fundamental claim of phonemic theory – that distinctive differences have a different linguistic status from non-distinctive ones. But phonemic theory goes beyond this primary claim: it posits that groups of mutually non-distinctive sounds are grouped together into categories, that is, the phonemes.

On the whole, linguists have found that speakers usually believe that two allophones of the same phoneme are the "same sound," despite the phonetic difference between them. Here are some examples.

The words *ten* and *Ted* have phonetically different vowels. In *ten*, the phoneme /ɛ/ occurs before a nasal sound, and is therefore eligible for a phonological rule which we can state as follows.

Vowel Nasalization
[+vowel] → [+nasal] / ___ [+nasal]
A vowel is realized as nasalized when it precedes a nasal consonant.

The application of the rule is shown below:

ten	*Ted*	
/tɛn/	/tɛd/	underlying forms
ɛ̃	—	Vowel Nasalization
[tɛ̃n]	[tɛd]	surface forms

In my experience, English speakers are quite willing to say that *ten* and *Ted* have "the same vowel." Indeed, the difference between the two sounds is felt to be subtle, and observable only with careful attention, for example, by greatly prolonging them.

For this phonetic difference, it is useful to compare the behavior of (for example) French speakers. Minimal pairs show that in French, nasal vowels are phonemically distinct from oral vowels:

[mɛ]	'but'	vs.	[mɛ̃]	'hand'
[tʁɛ]	'very'	vs.	[tʁɛ̃]	'train'

For French speakers, it is plain that [ɛ] and [ɛ̃] are different sounds. The crucial difference between a French speaker and an English speaker in this respect is the phonemic structure of the two languages: corresponding nasal and oral vowels in French count as different sounds because they are different phonemes; they count as the same sound in English because they are allophones of the same phoneme.

The same would hold true for the case of dental [t̪] and alveolar [t] in English and Bengali, mentioned above. Very few if any English speakers would sense that the [t̪] in *eighth* is not the "same sound" as the [t] in *eight*. Bengali speakers, however, have a strong sense that [t̪] and [t] are different.

We can conclude that, to a rough approximation, if two phones are allophones of the same phoneme, a speaker of the language in question will feel that they are the same sound.

However, we must supplement this conclusion with certain reservations about conscious awareness of sounds in general. It would appear that the natural state of humans is to be unaware that languages have speech sounds at all. What leads speakers to become consciously aware of their phonemes is typically the process of learning to read and write in an alphabetic system, since this focuses the learner's attention on the sounds that correspond to the letters.[2]

[2] A stronger position sometimes taken is that phonemic awareness is based *exclusively* on orthography. I'm skeptical of this position, in part because of my experience as a teacher: it is really quite easy to teach people to hear phonemic distinctions, such as /θ/ vs. /ð/, that are spelled inconsistently. Moreover, there are English dialects that have phonemes that have no distinctive spelling (for example, /ʌɪ/, in contrast with /aɪ/; or /ɛə/, in contrast with /æ/). In my elicitation experience, native speakers of these dialects have proven highly aware of these phonemes.

To state the original claim more accurately, then: *once speakers have been made aware* of the existence of speech sounds in their language, they will naturally tend to consider allophones of the same phoneme as counting as the same sound.

3.2.3 Foreign accents and transfer

The third fundamental claim made by phonemic theory is that speakers internalize rules that derive the various allophones in their appropriate environments. One of the best sources of evidence for the existence of such rules comes from second language acquisition, in particular, the behavior of speakers who are attempting to pronounce the sounds of a language new to them. The usual result, at least at first, is a rather poor imitation of the second language. Indeed, a foreign accent often persists even after years of practice with a second language.

Foreign accents are not the result of just "missing the mark" in random ways. To the contrary, careful inspection shows that the deviations between the goal and what is achieved are systematic; and can usually be attributed to the phonology, including the phonological rules, of one's native language. The phenomenon of mispronunciations in a second language in ways attributable to the phonology of the first language is called **transfer**.

To understand transfer, it helps to consider a phonology as specifying the set of *things that are pronounceable* in a given language. This set consists of the legal sequences of phonemes, realized as the appropriate allophones for their context. In other words, anything outside this (very large) set will necessarily involve one of three properties:

- It can be phonologically illegal (in a given language) because it contains an **illegal phoneme**. For example, in English any utterance containing the voiced uvular fricative [ʁ], a voiced aspirated stop, or a front rounded vowel is illegal.
- It can be phonologically illegal because it corresponds to an **illegal sequence** of phonemes (even where the phonemes themselves are legal). Thus, [bnɪk] consists of four English phonemes, arranged in an order which English phonology does not permit. (This point is treated in more detail below in §3.6 below.)
- It can be phonologically illegal because it corresponds to an **impossible distribution of allophones**. For example, [fil], with a non-velarized [l], is illegal in English, because English has an allophonic rule requiring the use of velarized [ɬ] in word-final position, as in the correct pronunciation [fiɬ] 'feel'.

If a word of a foreign language is phonologically illegal in English, for any of the above three reasons, it will typically not be pronounced correctly by English speakers, at least without practice.

Here are some examples. The German proper name *Gödel* is phonetically ['gøːdəl], with a long upper mid front rounded vowel. Many English dialects have no /øː/ phoneme, but do have the acoustically similar phoneme /ɚ/. Speakers of

these dialects tend to pronounce *Gödel* as *girdle*. French *thé* 'tea' is phonetically [t̪e]. Most English speakers speak a dialect in which the monophthong [e] does not occur; the diphthong [eɪ] occurs instead. These speakers must fight the tendency to substitute their own [eɪ] for French [e]; less conspicuously, they need to suppress their own native alveolar [t] and use dental [t̪] instead.

These are cases in which foreign accents arise from substituting native phonemes for phonetically similar foreign ones. We also get cases of substitution of native allophones for phonetically similar foreign sounds, through application of the phonological rules. For example, a typical English mispronunciation of French *thé* is [tʰeɪ], with an aspirated [tʰ], since that is the allophone of /t/ that occurs word-initially in English. Likewise, *Gödel* as pronounced by English speakers tends to receive the velarized [ɬ] allophone of /l/ (p. 25), which further mutilates it to ['gɝ·dəɬ].

The path from foreign source word to native rendition is sometimes a bit more crooked. French *tante* 'aunt' [t̪ɑ̃t] has a nasalized vowel alien to the English phoneme inventory. In an English accent, this usually comes out [tɑ̃n̆t], where [n̆] is a particularly short [n]. The derivation (that is, in English) here seems to be something like:

/tɑnt/	underlying form: choice of native phonemes	
ɑ̃	**Nasalization** (p. 50):	[+vowel] → [+nasal] / ___ [+nasal]
n̆	**Nasal Consonant Shortening**:	$\begin{bmatrix}+\text{consonant}\\+\text{nasal}\end{bmatrix} \rightarrow [+\text{short}] / \_\_\_ \begin{bmatrix}+\text{consonant}\\-\text{voice}\end{bmatrix}$
[tɑ̃n̆t]	surface form	

You can see that the derivation is rather "strategic": the insertion of /n/ disagrees sharply with the original French, yet this /n/ gives rise to an allophone, [ɑ̃], which considerably improves the matchup to the French vowel phoneme /ɑ̃/. Moreover, since nasal consonants are allophonically shortened before voiceless consonants in English, the offending /n/ ends up reduced in its perceptual salience. It seems to be the resemblance of the end result to the foreign word, and not some kind of step-by-step procedure, that guides the choice of phonemic accommodation.

To sum up, to a greater or lesser extent, learners of foreign languages are prisoners of their own phonologies. Lifelong experience leads them strongly to favor the legal phonological sequences of their native language, which may be defined as the sequences of allophones that are derived from legal sequences of native phonemes. The psychological reality of the constraints of the native phonology becomes blatant when one sees them determine the outcome of the native speaker's attempts to pronounce a second language.

Naturally, individuals differ greatly in their ability to overcome the transfer effect; in other words, to assimilate a novel phonology. In principle, explicit knowledge of the phonology of both native and learned language could be of help to foreign language learners in achieving a correct accent, and some foreign language textbooks do provide systematic training of this kind.

3.3 The Criterion of Phonetic Similarity

With the introduction of a criterion of psychological reality into phonological analysis, we find that in a number of cases, phonemicization cannot be done in purely mechanical fashion, as was done in the Maasai example of chapter 2. Here, we will consider some cases in which merely collecting and arranging the non-contrasting phonetic segments is insufficient.

A simple case arises in English, involving the sounds [h] and [ŋ]. [h] occurs at the beginnings of words and before stressed vowels, as in the examples below:

hill	['**h**ɪl]	*ahead*	[ə'**h**ɛd]
high	['**h**aɪ]	*prohibit*	[proʊ'**h**ɪbɪt]
how	['**h**aʊ]	*behold*	[bi'**h**old]
Horatio	[**h**ə'ɹeɪʃoʊ]	*rehearse*	[ɹi'**h**ɚs]

The sound [ŋ] occurs at the ends of words, before consonants, and (at least in the dialects we will consider) between vowels of which the second is stressless:

sing	['sɪŋ]	*sink*	['sɪŋk]	*singer*	['sɪŋɚ]
pang	['pæŋ]	*anger*	['æŋgɚ]	*Singapore*	['sɪŋəˌpɔɹ]
running	['ɹʌnɪŋ]	*hangs*	['hæŋz]	*dinghy*	['diŋi]
		Langley	['læŋli]	*hangar*	['hæŋɚ]

There are no cases of [h] occurring at ends of words, or before consonants, or between vowels of which the second is stressless. Likewise, there are no cases of [ŋ] occurring at the beginning of a word, or before a stressed vowel. Therefore, [h] and [ŋ] *do not contrast*. Given the phonological patterning of English, there is no way that they could distinguish words from each other, because they occur in entirely different contexts.

Since [h] and [ŋ] do not contrast, should we regard them as allophones of a single phoneme? The traditional answer of phonologists is no, for the following reason. When two sounds are allophones of the same phoneme, they will be felt by native speakers to be the same sound. This is plainly not the case for [ŋ] and [h]. It seems appropriate here to say that /h/ and /ŋ/ are separate phonemes of English, and that for accidental reasons (having to do with where it is legal for them to occur), they are unable to form contrasts. In other words, we reject the idea that phonemes can be established purely on distributional grounds; rather, if we are to posit that two sounds are allophones of a single phoneme, they must be related to each other phonetically in some way.

A similar example makes the same point. By reexamining the distributions of Maasai [p b t d k g] given on pp. 39 and 41, you should convince yourself of the following facts:

- [p] and [d] are in complementary distribution.
- [t] and [g] are in complementary distribution.
- [k] and [b] are in complementary distribution.

This is true because the voiced stops in Maasai always come after nasals, and the voiceless ones never do. We could imagine, then, a phonemic analysis of Maasai that grouped [p] and [d] together into the same phoneme, and similarly with [t]–[g] and [k]–[b]. Such an analysis indeed works on distributional grounds, but would almost certainly fail as a means of capturing the intuitions of the native speaker, who sensibly regards the phonetically similar [t] and [d] as being the same sound, and similarly with [p]/[b], [k]/[g].

At the level of phonological theory, it seems that we must impose some criterion of **phonetic similarity** on analyses: the allophones of a single phoneme should resemble one another to a particular degree. The case of Maasai is rather straightforward, since all we are trying to do is compare a sensible analysis, in which [b] is an allophone of /p/, [d] of /t/, and [g] of /k/, with a bizarre alternative that mixes up the places of articulation. By any criterion, the sensible analysis makes the allophones more mutually similar than the bizarre analysis does, and thus is to be preferred.

The case of /h/–/ŋ/, where we choose not to group two sounds into one phoneme despite their complementary distribution, is harder, because it raises the question of where to draw the line: what is the degree of dissimilarity beyond which the association of sounds into a single phoneme becomes impossible? [h] and [ŋ] are certainly far enough apart, but there are many other cases where the outcome is not so clear. It is possible that during language change, new phonemes might be created when two allophones drift apart too far to count any more as variants of the same basic linguistic unit.

3.4 Other Problems in Phonemicization

Defining the phonetic similarity criterion rigorously is by no means the only conundrum in traditional phonemicization; here are some others.

3.4.1 Contour segments and the segment/sequence problem

Sounds like diphthongs ([a͡i]), affricates ([t͡ʃ]), and prenasalized stops ([m͡b]) are often called **contour segments**: they have two phonetic qualities in sequence, but are often treated phonologically as a single sound. The recognition of the contour segments is often an analytic difficulty faced in phonemicization. For example, given a sequence like [ai] in the data, we need to decide whether it should be treated

as a diphthong (that is, as a single phonemic unit), or as a sequence of /a/ + /i/. The same issue arises for [tʃ] (is it the affricate [t͡ʃ], or is it /t/ + /ʃ/?) and for [mb, nd, ŋg] (prenasalized stops or *nasal* + *stop* sequences)? This analytical issue might be called the **segment/sequence problem**.

The problem is easy to solve if there is an actual contrast between segment and sequence. For instance, the analysis of [t͡ʃ] as an affricate in Polish is uncontroversial, because this sound contrasts with the stop + fricative sequence [tʃ]. The following minimal pair illustrate this:

[tʃɨ]	*trzy*	'three'
[t͡ʃɨ]	*czy*	'if, whether'

[tʃ] and [t͡ʃ] are phonetically different; in particular, [tʃ] is noticeably longer than [t͡ʃ]. The affricate [t͡ʃ] must be analyzed as a single segmental unit in Polish, since otherwise we could not express the contrast between monosegmental /t͡ʃ/ and bisegmental /tʃ/. Similarly, in other languages diphthongs contrast with two-vowel sequences (English *boing* [bɔ͡ɪŋ] vs. *sawing* [sɔɪŋ]), and prenasalized stops contrast (for example in Sinhala) with nasal + stop sequences. In Turkish, long vowels contrast with identical vowel sequences, as in [daː] 'mountain-nom. sg.' vs. [da.a] 'mountain-dat. sg.', which leads to the clear conclusion that the long vowels must be single phonemic units in Turkish.

Sometimes phonologists choose between segments and sequence on the basis of the overall pattern of the language. Suppose that a language has five monophthongs /i, e, a, o, u/, and that moreover any one of the 25 logical possibilities for putting any two of these vowels together occurs in the language; that is, we observe:

[iː]	[ie]	[ia]	[io]	[iu]
[ei]	[eː]	[ea]	[eo]	[eu]
[ai]	[ae]	[aː]	[ao]	[au]
[oi]	[oe]	[oa]	[oː]	[ou]
[ui]	[ue]	[ua]	[uo]	[uː]

In this case, we would be sensible to opt for a sequence analysis, in which these putative "diphthongs" are not phonemes, but merely sequences of the independently existing vowel phonemes /i, e, a, o, u/, as follows:

/ii/	/ie/	/ia/	/io/	/iu/
/ei/	/ee/	/ea/	/eo/	/eu/
/ai/	/ae/	/aa/	/ao/	/au/
/oi/	/oe/	/oa/	/oo/	/ou/
/ui/	/ue/	/ua/	/uo/	/uu/

The two criteria just given do not always suffice to determine an analysis. The apparent diphthongs of Mandarin Chinese are a classical example; they have been

treated by some analysts as monophonematic diphthongs and by others as sequences. To consider just a subset of the Mandarin problem, observe that Mandarin has the following sounds: [ə], [i], [u], [ei], and [ou]. [e] and [o] never occur alone, but only as part of the diphthongs [ei] and [ou]. One possible phonemicization is the following:

/ə/	/əi/	/əu/	/i/	/u/	underlying forms
—	ei	—	—	—	Vowel Assimilation I: ə → e / ___ i
—	—	ou	—	—	Vowel Assimilation II: ə → o / ___ u
[ə]	[ei]	[ou]	[i]	[u]	surface forms

The appeal of this analysis is that we can get by with just three phonemes (/ə/, /i/, /u/) to derive five sounds. Moreover, the rules make sense as **assimilation** rules; the vowel /ə/ is assimilated to [i] or [u], become phonetically more similar to its neighbor; and assimilation is a very common process in phonology. However, not all phonologists would necessarily agree that this is an iron-clad argument – in principle, we want to know not just a convenient and elegant way to symbolize Mandarin sounds, but the way that actually is found in the internalized grammars of native speakers. In general, the issue of how to segment the speech stream into its phonemes is an unsettled one in phonology.

3.4.2 Borrowed sounds

We live in a time in which the phonology of a great number of the world's languages is in flux, as a result of new borrowed sounds. When a borrowed sound is used for the very first time by a single speaker, it cannot count as a phoneme of the language. But with time, borrowed words come to be used by larger numbers of speakers. Eventually, they are felt by native speakers to be an integral part of the language. One example is English /v/, which centuries ago was a borrowed phoneme (see §11.7) but is now fully integrated into the English phoneme inventory.

The difficulty for phonological analysis is that the process is gradual. Contemporary Japanese provides an example. In Japanese as it was spoken not long ago, the sound [ɸ] (voiceless bilabial fricative) was plainly an allophone of /h/. It occurred only in the environment / ___ u, and was in complementary distribution with [h], which occurred in most other environments and thus was the elsewhere allophone.

As Japanese has evolved under the influence of English and other foreign languages, [ɸ] has extended its usage: it is the usual way to approximate a foreign [f] sound. If we are willing to include recently borrowed words in our data set for phonemicization, then we can now find many near-minimal pairs for [ɸ] vs. [h], as follows:

before /a/:	[ɸaito] 'fight', [ɸaŋ] 'fan'	(vs. [**h**aiku] 'type of poetry')
before /e/:	[ɸesutibaɾu] 'festival', [ɸeɾuto] 'felt'	(vs. [**h**ema] 'blunder')
before /i/:	[suɸiŋkusu] 'sphinx', [ɸiɾumu] 'film'	(see footnote[3])
before /o/:	[ʃiɸoŋ] 'chiffon', [ɸoːku] 'fork'	(vs. [**h**oŋ] 'book')

To the extent that we can consider the words in the left column to be authentic words in the vocabularies of innovating speakers, we must say that the dialect spoken by these speakers has acquired a new phoneme, having promoted [ɸ] from allophone to phoneme status.

But the crucial question is whether such words really *can* be considered authentic words of contemporary Japanese. Different speakers feel differently about whether the words above are normal, everyday words. Speakers of the more conservative variety of Japanese consider them to be exotic imports. This may be compared with the way English speakers regard words and phrases like [ˌfæ̃ də siˈɛklə] 'end of the century', ['bɑ**x**] 'Bach', and [t͡**s**uˈnɑmi] 'wave caused by an earthquake',[4] taken from French, German, and Japanese respectively.

That there really is a continuum is suggested by the fact that there are some sounds that are less far along in the process of becoming Japanese phonemes. According to my reference source (see below), younger, cosmopolitan speakers of Japanese have an emergent phonemic distinction between /d/ and /d͡z/; for all other speakers [d͡z] is merely the allophone of /d/ that occurs before /u/. For these younger speakers, in certain speaking styles, it is possible to say words like [**duː** itto juaseɾuɸu] 'do-it-yourself' and [**duː**-wappu] 'doo-wop (music)'.

In the end, we must recognize that a phonemic analysis is a phonemic analysis of a **particular stratum** of the vocabulary. Such strata can include the core vocabulary, the vocabulary as it includes mildly non-native words, or even the vocabulary as amplified by words still recognized as quite foreign. The blurriness of the phonemic analysis arises from the blurriness of the concept "member of the language's vocabulary." Careful presentation of a phonemic analysis must therefore specify the vocabulary strata on which the analysis rests.

3.5 Free Variation

Languages are permeated with variation: we frequently say the same thing in different ways. In phonology, free variation takes two forms. One is the phenomenon of **phonological doublets**, in which one word happens to have two different phonemic forms. For instance, in many people's speech, the word *envelope* can be pronounced as either ['ɛnvəˌloʊp] or ['ɑnvəˌloʊp]; *economics* as either

[3] Before /i/, /h/ has the allophone [ç] (voiceless palatal fricative), as in /himo/ [çimo] 'string'.

[4] These are how I usually pronounce these words when using them in English. These pronunciations are not the same as those used in the source languages.

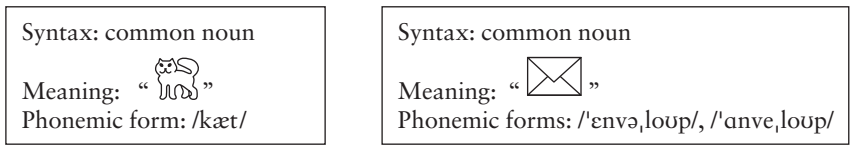


Figure 3.1 Lexical entries for “cat” and “envelope”

[ˌikəˈnɑmɪks] or [ˌɛkəˈnɑmɪks], *deity* as either [ˈdiəɾi] or [ˈde͡iəɾi]. This does *not* refer to instances in which different people say certain words differently; rather, a doublet is a case where one and the same person uses both variants.

The usual treatment of phonological doublets posits that in the **lexicon** (the mental store of words in the mind/brain), they have just one listing for their syntactic properties and meaning, but more than one phonemic representation, as shown in figure 3.1.

The other kind of variation in phonology is when a single phonemic representation gives rise to more than one phonetic form; this is called **free variation**. Here is one example found in the speech of many Americans. In the dialect in question, the vowel phoneme /æ/ has a diphthongal allophone I will transcribe as [ɛ̃ə̃]. Some data on the distribution of [ɛ̃ə̃] vs. [æ] are given below:

	[æ]			[ɛ̃ə̃]	
lap	/læp/	[læp]	*man*	/mæn/	[mɛ̃ə̃n], [mæ̃n]
pal	/pæl/	[pæl]	*Spanish*	/spænɪʃ/	[spɛ̃ə̃nɪʃ], [spæ̃nɪʃ]
pack	/pæk/	[pæk]	*dance*	/dæns/	[dɛ̃ə̃ns], [dæ̃ns]
lab	/læb/	[læb]	*flannel*	/ˈflænəl/	[ˈflɛ̃ə̃nəl], [ˈflæ̃nəl]

To summarize the pattern: if an /n/ follows /æ/, then there are two outputs, one with [ɛ̃ə̃] and one with [æ̃]. Otherwise, the observed allophone is [æ]. This [ɛ̃ə̃]–[æ̃] pattern is systematic; it holds not just for these four words, but for any word in this dialect in which /æ/ precedes /n/. Indeed, even if we make up a completely new word (“Hello, my name is *Thran*”), it too would behave this way: [θɹɛ̃ə̃n], [θɹæ̃n].

In analyzing the data, we should first dispose of the distribution of nasality. The nasalization seen on both [ɛ̃ə̃] and [æ̃] is plainly the consequence of Vowel Nasalization (p. 50). More crucial is the free variation between the monophthongal and diphthongal allophones. These cannot be phonological doublets, in the sense given above, because they are part of a systematic pattern rather than being idiosyncratic. We need to express the variation with a rule.

An appropriate analysis, then, would be as follows. We set up /æ/ as the basic form of the phoneme, and include the following rule.

/æ/ Diphthongization (preliminary)
æ → ɛə / ___ n
The phoneme /æ/ *is realized as* [ɛə] *when it precedes* /n/.

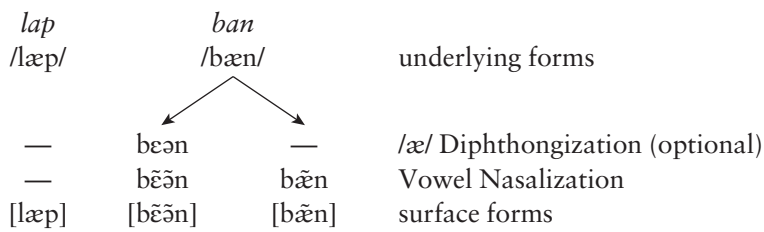
Together with Vowel Nasalization, /æ/ Diphthongization suffices to derive the [ɛ̃ə̃] variants, as illustrated below.

ban:	/bæn/	underlying form
	bɛən	/æ/ Diphthongization
	bɛ̃ə̃n	Vowel Nasalization
	[bɛ̃ə̃n]	surface form

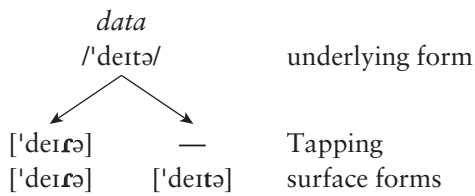
We can derive the non-diphthongized variants as well, provided that we indicate that /æ/ Diphthongization **applies optionally**; that is, on any given speaking occasion, the speaker may apply the rule or not. I will designate optional rules simply by including the word "optional" in parentheses after the statement of the rule:

/æ/ Diphthongization (revised)
æ → ɛə / ___ n (optional)
The phoneme /æ/ may be realized as [ɛə] *when it precedes* /n/.

We can explicitly show the effects of optional rules with **branching derivations**, which include arrows to indicate what happens when an optional rule does or does not apply. The derivation for *ban* below is a branching derivation.



Another optional rule of English, at least for some speakers, is the Tapping rule discussed on p. 32. For such speakers, the derivation of *data*, given earlier, should actually be a branching derivation:



As with /æ/ Diphthongization, the phenomenon occurs across the board, and is not specific to the word *data*; similar variation is found for *outer* ['aʊɾɚ]/['aʊtɚ], *attic* ['æɾɪk]/['ætɪk], *catapult* ['kæɾəpʊlt]/['kætəpʊlt], and so on.

Not all phonological rules are optional. Vowel Nasalization, for instance, is obligatory: while [bɛ̃ə̃n] and [bæ̃n] are both options for *ban* (in the relevant dialect), *[bɛən] and *[bæn], with oral vowels, are impossible. (I follow normal practice

in linguistics in using asterisks to designate impossible forms.) Similarly, the rule that assimilates alveolar /n/ to dental [n̪] before /θ/ (p. 49) is obligatory: to say the word *tenth* as [tɛnθ], with a true alveolar [n], is difficult and artificial.

The same rule can be optional in one dialect of a language and obligatory in a closely related dialect; this is true, for instance, of both Tapping and /æ/ Diphthongization.

3.5.1 Optional rules and speech style

When a language has an optional rule, it is often the case that the choice of whether to apply the rule or not is determined by the style of speech. A typical pattern is that in solemn or careful speech, application of an optional rule is suppressed, whereas in casual or rapid speech, it will apply. This can be seen clearly if you pick a sequence of words that allows for the application of more than one optional rule. For example, the phrase *tan attic* is phonemically /ˈtæn ˈætɪk/. It can undergo Tapping and [æ] Diphthongization (along with obligatory Vowel Nasalization) to become [ˈtɛ̃ɚ̃n ˈæɾɪk], a casual pronunciation. Alternatively, Tapping and [æ] Diphthongization could be suppressed, yielding [ˈtæ̃n ˈætɪk], a solemn and careful pronunciation. There could also be intermediate variants, in which only one of the optional rules is suppressed.

In most languages and dialects, speakers command a broad range of speaking styles, of which they may be only dimly aware. This variation is amenable to phonological analysis: typically, there are particular optional phonological rules that apply with greater frequency (or derive more dramatically deviating outputs) when the speaker is in a more casual social context. Speakers unconsciously tune their phonological style so as to produce appropriate speech behavior for each context.

The word "appropriate" in the previous paragraph deserves clarification. In literate societies, people are often told how they *should* speak; that is, there are many **normative** beliefs about language, which include beliefs about pronunciation. The sense of "appropriate" here is descriptive, not normative. While there exist contexts where informal speaking styles are felt to be inappropriate, there are also contexts where formal styles are likewise perceived as inappropriate – that is, as pretentious or pedantic.

3.5.2 Variation and elicitation

Free variation can pose problems for the linguist who is eliciting new data from a native speaker consultant. For one thing, the linguist is seldom in a good position to get the speaker to say each variant – usually the speaker will say the one that is appropriate for the social context (namely, elicitation), and the other variant(s) will go undetected.

The missing variant will normally be the one characteristic of informal speech. The reason is that most consultants construe elicitation, where attention is focused

on their speech, to be a rather formal speaking context. Simply requesting the consultant to speak casually may not help much, since speakers often have little conscious control over their speaking style.

The problem of "missing free variants" often arises when a student attempts to hear from a native speaker the application of a rule that she has read about in a reference source. If the rule applies only in more casual speech, and the consultant adopts a formal speaking style, then the researcher may wrongly draw the conclusion that the rule simply doesn't exist in the consultant's speech.

Here is some practical advice on how to elicit informal variants from a consultant. First, one can at least make clear to your consultant that you are interested in casual as well as formal speech. Second, it often pays to turn off any recording equipment and rely (at least temporarily) on one's ability to transcribe on the spot, since recording equipment often can intimidate the consultant. Third, it usually helps to elicit words in **connected speech** rather than in isolation (say "Could you use that word in a sentence?"); connected speech almost always produces more rule application than citation forms. Lastly, in a pinch it sometimes helps to ask a consultant to say the same thing many times; often boredom will set in and the consultant will speak more casually.

The ultimate method for investigating casual speech is to abandon elicitation entirely, and let the consultant speak spontaneously. This is the central research method of **sociolinguistics**, a field which has developed methods for studying casual speech systematically (see Further reading, at the end of this chapter). Recorded conversations among family members or friends[5] will often reveal optional phonological rules that could never have been detected in elicitation. Of course, the disadvantage of sociolinguistic methods is that they require a great deal of work to get the data, since one must wait until the data of the type one is looking for turn up by accident.

When one's goal is to elicit *formal* speech, the task is of course much easier. Elicitation tactics that are likely to evoke formal speech include: (a) asking the consultant to speak "correctly"; (b) use of recording equipment; (c) asking the speaker to pronounce individual words, especially minimal pairs; and (d) asking the consultant to make a recording in a soundproof booth, if one is available.

3.6 Contextually Limited Contrasts and Phonotactics

Phonemic contrast is often not an across-the-board matter, but is confined to particular contexts. For instance, in Toba Batak (Austronesian, Sumatra), there

[5] It helps to let the recording equipment run for quite a while before transcribing; this permits the consultants to become more used to its presence.

is a general contrast between voiced and voiceless stops and affricates, as the following near-minimal pairs attest:

[**p**inoppar]	'descendant'
[**b**iaŋ]	'dog'
[**d**ukkar]	'let out'
[**t**uak]	'palm wine'
[**k**orea]	'Korea'[6]
[**g**arut]	(name of town in Indonesia)

Many words of Toba Batak also end in voiceless stops:

[sukku**p**]	'adequate'
[hoto**p**]	'fast'
[doho**t**]	'with'
[sura**t**]	'letter'
[rappo**k**]	'steal'
[hala**k**]	'man'

However, no word in the language ever ends in a voiced stop: hypothetical words like *[sukku**b**], *[doho**d**], or *[rappo**g**] sound "un-Batak" to native speakers.

Thus, we have a phonological contrast of voicing, but it is a **contextually limited contrast**. A full description of Toba Batak must include a characterization not just of the contrasting phonemes, but also a characterization of where the contrast is allowed.

3.6.1 Analyzing phonotactics and contextually limited contrast

Two formal approaches have been taken to contextually restricted contrast. In one, we write rules that would have the effect of eliminating the contrast. For the Toba Batak case given above, a suitable rule would be the following:

Final Devoicing (Toba Batak)
[+stop] → [–voice] / ___]$_{word}$
Stops are devoiced at the end of a word.

This approach may seem slightly counterintuitive, since one wonders: what are the forms to which the rule applies? There is no reason to set up any underlying forms in Toba Batak that would qualify. The idea behind positing such a rule is

[6] A caution: /k/ is a borrowed phoneme in Toba Batak, so that this (unlike /p/–/b/ and /t/–/d/) is a marginal contrast.

to say, "even if Toba Batak *did* have final voiced stops in its underlying forms, they would be pronounced as voiceless in surface forms." The result is in fact a correct prediction, as no Toba Batak word can end with a voiced stop.

A different approach to contextually limited contrast posits that phonological theory involves not just rules but also **constraints**. A constraint is a formal characterization of a structure that is illegal in a particular language. In the constraint below, the asterisk may be read "is illegal" or "is ill-formed."

Constraint against Final Voiced Stops (Toba Batak)

$$*\begin{bmatrix} +\text{stop} \\ +\text{voice} \end{bmatrix} / \_\_\_]_{\text{word}}$$

It is illegal to have a voiced stop in word-final position.

Such constraints are sometimes called **phonotactic constraints**, "phonotactics" being a general term for the principles (however stated) of phonological well-formedness in a particular language. Phonologists debate what are the roles and relative importance of rules and constraints in phonology. Some theories make use only of rules, some use both, and some theories use only constraints.

3.6.2 Contrast with zero

The notion of phonological contrast can be broadened to include **contrast with zero**. For instance, English allows contrasts like *tax* [tæks] vs. *tack* [tæk], where the [s] of *tax* is said to be in contrast with zero. The following diagram illustrates this; ∅ is the symbol representing the null string.

t æ k s
t æ k ∅

Contrast with zero can also be contextually limited. Thus, for instance, Toba Batak has no contrasts like the one just given, because it never permits two consonants to occur at the end of a word. Consonants do contrast with zero in other environments of Toba Batak, for example in the context / V ___ $]_{\text{word}}$. One of many examples would be [laŋa] 'empty' vs. [laŋan] 'pale'. In fact, Toba Batak falls into a very widespread phonological pattern whereby consonants may contrast with zero only when they are adjacent to a vowel.

To ban the CC vs. C contrast in final position, we formulate either a rule or a constraint, as shown below. Note that "X → ∅" is the usual notation in phonology for deletion.

Rule: Cluster Simplification

$C \rightarrow \varnothing / C \_\_\_]_{\text{word}}$

Delete a word-final consonant if a consonant precedes.

Constraint: Ban on final clusters
*CC]$_{word}$

In comparing these two approaches, we see one possible objection to the rule-based theory: it often forces us to make arbitrary analytic decisions. In particular, given the data we have, there seems to be no reason to delete the second consonant rather than the first (C → ∅ / C ___]$_{word}$); both rules would suffice to enforce the one-consonant limit.[7] Often, there is further evidence from the language that tells us which rule is correct; see chapter 6. In addition, one should bear in mind that it is perhaps not so bad to have two possible analyses available, when both of them happen to work.

Concluding up to this point, we have now considered most, though not all, of the basic goals of phonological analysis. In analyzing a language, we seek first to isolate its inventory of phonemes. The allophonic variation of phonemes, both contextual (chapter 2) and free (§3.5), must be characterized with appropriate phonological rules. Lastly, the limitations on contrast, both between phonemes and between phonemes and zero, must be characterized with rules or constraints. There remains one further (large) area, phonological alternation, to which we will turn in chapter 6. Since alternation is dependent on morphology, chapter 5 gives an overview of this topic.

Exercises

1 Psychological reality of the phoneme in Ilokano

In the following (partly true, partly fictional) dialogues, BH is me. MA is a UCLA undergraduate. MA had studied some linguistics, including phonetics, and was collaborating with me on a research project on Ilokano (Austronesian, Northern Philippines), which is her native language.

Ilokano has no minimal pairs for [ʔ] vs. [t]. [ʔ] is derived from /t/, by an optional rule of **/t/ Glottalization**: t → ʔ / ___ C. The dialogue is from an elicitation session that focused in part on /t/ Glottalization.

BH$_1$: "egg"
MA$_1$: "[it'log]"
BH$_2$: "Say it three times, please."
MA$_2$: "[it'log], [iʔ'log], [iʔ'log]"
BH$_3$: "men"

[7] More precisely, they suffice if applied iteratively (see §14.4) to make sure that clusters of any length are reduced to just one.

MA_3: "[tat'tao] . . . [taʔ'tao]"
BH_4: "to fall"
MA_4: "[mat'twaŋ]"
BH_5: "Would [maʔ'twaŋ] also be ok?"
MA_5: "Sure, [maʔ'twaŋ] is fine."
BH_6: "the egg of the chicken"
MA_6: "[ti it'log ti ma'nok]"
BH_7: "Would [ti iʔ'log ti ma'nok] be ok?"
MA_7: "Yes."
BH_8: "What about [ʔi 'iʔlog ʔi ma'nok]?"
MA_8: "No, completely impossible."

Questions:

a. Why did BH request repetition at BH_2?
b. What explains MA's reply in MA_8?

Here is another dialogue, this time MA eliciting from her mother, JA:[8]

MA_1: "to smell like candy"
JA_1: "[agat'dulse] . . . [agaʔ'dulse]"
MA_2: "egg"
JA_2: "[it'log]"
MA_3: "Could you also say [iʔ'log]?"
JA_3: "Of course; I just *did*: [it'log]".

c. What do we learn about JA's idiolect[9] from JA_1?
d. Explain why JA said what she did at JA_3, making use of the material in this chapter about conscious awareness of speech sounds. You should assume that listening conditions were excellent and that JA's hearing is fine.
e. Compare JA_3 with MA_5 in the previous dialogue. Why are they different?

2 *Final stop clusters in English*

English words can end in two stops, but only if the second is alveolar:

concept	['kɑnsɛ**pt**]	contact	['kɑntæ**kt**]
jumped	[dʒʌm**pt**]	milked	[mɪl**kt**]
rubbed	[ɹʌ**bd**]	bagged	[bæ**gd**]

Thus, there are no words in English like *['kɑnsɛ**tp**], *['kɑntæ**tk**], *[mɪl**kp**], *[bæ**dg**], or *[ɹʌ**db**]. Speakers of English immediately recognize such hypothetical words as

[8] MA and JA are fluent bilinguals. They tend to use English in discussions of schoolwork and Ilokano for domestic topics.
[9] For "idiolect" see fn. 11, p. 35.

ill-formed, and often regard them as hard to pronounce. Following the discussion in §3.6 above, give two analyses, one based on rules, the other on constraints.

3 Contextually limited contrast in Japanese

In most dialects of Japanese, there are several nasal consonants: [m, n, ŋ, w̃, j̃] (the latter two are nasalized glides). However, not all of these are phonemic, and the phonemic distinctions are limited to certain contexts.

Examine the Japanese data given below.

a. There is a phonemic contrast for place of articulation in Japanese nasals. Find this contrast and give three minimal pairs to justify it.
b. Collect local environments (§2.10.3) for the sounds [m, n, ŋ, j̃, w̃].
c. The phonemic contrast you located in (a) occurs in a certain position. State in words what this position is.
d. In all other contexts, there is no contrast for place of articulation in nasals, and the place of articulation of nasals is predictable. State phonological rules that determine the place of articulation for these other contexts. If you cannot think of a formalism for a particular rule, either invent one or state the rule as clearly as possible in words.

1	[mi]	‘body’	24	[sɯpeiŋ]	‘Spain’
2	[me]	‘eye’	25	[deŋ]	‘palace’
3	[ma]	‘interval’	26	[nihomppoi]	‘Japanesey’
4	[muʃi]	‘neglect’	27	[zambuɾi]	‘with a splash’
5	[matte]	‘wait’	28	[ʃimmiɾi]	‘quiet sympathy’
6	[mo]	‘also’	29	[hommo]	‘book-too’
7	[ni]	‘two’	30	[aɾimasentte]	‘that there isn’t’
8	[na]	‘name’	31	[kentoː]	‘examination’
9	[no]	‘field’	32	[honda]	‘brand of automobile’
10	[nuʃi]	‘owner’	33	[jonda]	‘read-past’
11	[hema]	‘blunder’	34	[onna]	‘woman’
12	[ʃima]	‘island’	35	[kaɾendaː]	‘calendar’
13	[ɕimo]	‘string’	36	[hoj̃ja]	‘bookstore’
14	[zama]	‘state’	37	[boj̃jaɾi]	‘vacantly’
15	[jomu]	‘read-nonpast’	38	[zaŋki]	‘remaining time’
16	[oːkami]	‘wolf’	39	[ɾondoŋkko]	‘Londoner’
17	[samui]	‘cold’	40	[nippoŋgiŋkoː]	‘Bank of Japan’
18	[oneːsaŋ]	‘older sister’	41	[oŋgaku]	‘music’
19	[katana]	‘sword’	42	[geŋgogaku]	‘linguistics’
20	[tanuki]	‘raccoon dog’	43	[jaw̃waɾi]	‘softly’
21	[hoŋ]	‘book’	44	[how̃wa]	‘book-topic’
22	[paŋ]	‘bread’	45	[ikebana]	‘flower arranging’
23	[ʃizuŋ]	‘season’	46	[ɸud͡ʒijama]	‘Mount Fuji’

Further reading

Phonemic principle in alphabet design: Joseph E. Grimes and Raymond G. Gordon, Jr., "Design of new orthographies," in James F. Kavanagh and Richard L. Venezky, eds., *Orthography, Reading, and Dyslexia* (1980, University Park Press).

Effect of native phonology on the audibility of phonetic distinctions: a literature summary with references is given on pp. 105–6 of Sharon Peperkamp (2004) "Lexical exceptions in stress systems: arguments from early language acquisition and adult speech perception," *Language* 80: 98–126 [www.ehess.fr/centres/lscp/persons/peperkamp/Language.pdf]. The native-phonology effect is found even in infants under one year of age, as was originally shown by Janet Werker and Richard Tees (1984) "Cross-language speech perception: evidence for perceptual reorganization during the first year of life," *Infant Behavior and Development* 7: 49–63. For further references on phonemic perception in infants, see pp. 99–100 of the Peperkamp article.

Experimental evidence bearing on whether speakers hear two allophones of the same phoneme as "the same sound": Jeri J. Jaeger (1980) "Testing the psychological reality of phonemes," *Language and Speech* 23: 233–53; Bruce Derwing, Terrance M. Nearey, and Maureen L. Dow (1986) "On the phoneme as the unit of the 'second articulation,'" *Phonology Yearbook* 3: 45–70.

I know of little psycholinguistic work that has been done on how native speakers tacitly decide whether to treat phonetic sequences as contour segments or as phoneme sequences (§3.4.1). A paper that at least presents a method that could be useful in future work is Linnea C. Ehri and Lee S. Wilce (1980) "The effect of orthography on readers' conceptualizations of the phonemic structure of words," *Applied Psycholinguistics* 1: 371–85.

Sociolinguistics and casual speech data: work of William Labov, particularly *Sociolinguistic Patterns* (1972, University of Pennsylvania Press), *Language in the Inner City* (1972, University of Pennsylvania Press), *Principles of Linguistic Change*, Vols. 1 and 2. (1994, 2001, Blackwell). Peter Trudgill's *Sociolinguistics: An Introduction to Language and Society* (4th ed., 2001, Penguin) is a good short introduction.

Rules and constraints in phonology: two current textbooks cover Optimality Theory, a constraints-only approach. See René Kager (1999) *Optimality Theory* (Cambridge University Press) and John McCarthy (2001) *A Thematic Guide to Optimality Theory* (Cambridge University Press).

Polish affricates: Maria Zagorska Brooks (1965) "On Polish affricates," *Word* 20: 207–10. Borrowed phonemes in Japanese: Timothy Vance, *An Introduction to Japanese Phonology* (1987, State University of New York Press); Junko Ito and Armin Mester "Japanese Phonology," in *Handbook of Phonology*, John Goldsmith, ed. (1995, Blackwell). Turkish [daː] vs. [da.a]: G. N. Clements

and S. J. Keyser, *CV Phonology* (1983, MIT Press). Mandarin diphthongs: the analysis given is modeled on Lawton Hartman III (1944) "The segmental phonemes of the Peiping dialect," *Language* 20: 28–42. Final Devoicing in Toba Batak: P. W. J. Nababan, *A Grammar of Toba-Batak* and W. Keith Percival, *A Grammar of the Urbanised Toba Batak of Medan* (both 1981, Australian National University).