7 Segmental Phonology

Phonology frees us from a kind of nightmare which had weighed upon us.
Antoine Meillet, quoted by Jakobson in *Six Lectures on Sound and Meaning*

1. Introduction

In the next two chapters we present the basic phonological components that one finds in human languages, and we discuss various ways in which the structure of these components can be elucidated in the language spoken by your informants. We first consider phonological processes that apply to individual sound segments, and in chapter 8 we turn to phenomena that affect larger phonological domains such as syllables and words.

In order to carry out the basic phonological research that is described in the next two chapters, it is necessary to be familiar with a few fundamental concepts of phonological theory. For our purposes, the most basic of these concepts is that words have two different forms: the one that we store in our memory, and the one that we pronounce. Take, for example, the English negative morpheme[1] that we find in words such as *inept*, *indecent*, and *incapable*. Let us assume that we store this negative morpheme in our brains as */ɪn-/*. How then do we account for the fact that this morpheme is actually pronounced in several different ways? In *impossible*, for example, it is pronounced as [*ɪm-*], in *incapable* it is pronounced as [*iŋ-*], and in *inept* it is pronounced as [*ɪn-*]. The answer to this question is that the forms that we store in our brains, or *underlying representations*, can be affected by various types of *rules*, which cause the actual pronunciations of individual words—*surface representations*—to differ from their underlying representations.

In the case of *impossible*, for example, the underlying form is */ɪn-pasɪbəl/*. This undergoes a rule that requires */n/* to have the same place of articulation as a following consonant. The result of this rule is that the *n* assimilates to the labial place of articulation of the following *p*, thereby becoming *m*. The surface representation that ends up being pronounced therefore is [*ɪmpʰasɪbəl*] (we discuss in the next paragraph the rule that aspirates the *p*). The derivation of the word *impossible* can be schematized as follows:

underlying representation	*/ɪn-pasɪbəl/*
	↓
nasal place assimilation	*ɪmpasɪbəl*
	↓
aspiration	*ɪmpʰasɪbəl*
	↓
surface representation	[*ɪmpʰasɪbəl*]

Rules that apply to individual segments, which are the sort we will focus on in this chapter, have the effect of producing different surface pronunciations of single underlying segments. The underlying forms of these segments are called *phonemes*; the different surface forms in which a given phoneme appears are called *allophones*. The English phoneme */t/*, for example,

---

[1] Loosely speaking, a morpheme is the minimal unit of meaning in a language. In English, for example, *dogs* contains two morphemes: *dog*, which means "a domesticated carnivorous animal, *canis familiaris*", and *-s*, which means "plural". For more details concerning morphemes, see chapters 9 and 10.

has several different allophones:

- plain voiceless [*t*] in *stop*, *mistake*, etc.
- aspirated [*tʰ*] in *top*, *pretend*, etc.
- flapped [*ɾ*] in *later*, *party*, etc.
- glottalized and unreleased [*ʔt̚*] in (the American pronunciation of) *hat*, *Atlantic*, etc.
- palatal affricate [*ʧ*] in *tree*, *betray*, etc.[19]

In this chapter we will be interested in the questions of how one identifies the allophones that are used in a given language, and how we can figure out the rules and constraints that produce these allophones from the inventory of phonemes in the language. These ultimately reduce to the same question, since identifying allophones goes hand in hand with identifying rules and constraints, and vice versa. In what follows, we break our consideration of this question into two parts, the first dealing with rules, and the second with constraints.

2. How to Identify Phonological Rules

In order to identify rules in the language you are studying, you will first need to find alternations between different segments. These alternations can be either within a paradigm (as in the case of the vowel in Spanish *tengo* 'I have' vs. *tienes* 'you have'), within a single morpheme (as with English *in-*), or within a single segment in different environments (as with the allophones of English /*t*/ mentioned earlier). Identifying these alternations is essentially the same task as identifying the allophones in the language you are studying.

There are various ways of going about this task. The easiest way is to draw your information from existing grammars or dictionaries; if tools of this type are available for your language, you should make sure to consult them. However, these rarely contain all of the information you will need to construct a reasonable model of the phonological component of your informant's grammar: published descriptions tend to be incomplete, your informant's dialect may well differ from the one described in the dictionary or grammar, and so on. You therefore must check all published descriptions of the language against your own field notes. Sometimes, simply asking your informants directly whether or not their dialect contains a phonological rule that you have read about can be very effective.

But what should you do when there are no relevant materials written about your language? In this case, there are two ways to investigate the phonological system. If you have the right kind of informants, you can of course ask them if they have noticed any alternations in their own language. This can produce surprisingly useful results, but often does not. It is important to remember that speakers typically are not aware of fast speech and allophonic rules (that is, rules that produce sounds that are not part of the phonemic inventory of the language). For example, if you were to ask speakers of English if their language contains any alternations between different sounds, they might well think of [*k*] ~ [*s*] in *electric* : *electricity*, [*i*] ~ [*ɛ*] in *serene* : *serenity*, and so on. On the other hand, they would almost certainly be unaware of rules such as *s*-retroflexion (*street* → [*ʃtʃrit*]), aspiration (*pin* → [*pʰɪn*]), devoicing (*bun* → [*pʌn*]), and palatalization (*keep* → [*cʰip*]). In fact, most individuals would probably also fail to notice more obvious phenomena such as the *t*-insertion that many speakers have in words like *else* [*ɛlts*] and *false* [*falts*].

You should also be aware of the fact that literate speakers are often influenced heavily

[19] This allophone of /*t*/ is actually produced with a tongue position intermediate between that of [*t*] and of [*ʧ*], but we have used [*ʧ*] here for typographical convenience.

by spelling. Surprisingly often, speakers will try to alter their pronunciation to match what they know to be the spelling of a given word, particularly when they are speaking carefully. Even when they do not alter their pronunciation, informants are generally reluctant to acknowledge that their pronunciation differs from the sequence of letters in the official orthography. For example, speakers of English will typically insist that the *p*'s in *pit* and *spit* are the same, even though the former is aspirated and the latter is not. Moreover, speakers of dialects which possess sounds not found in the literary dialect will deny that they pronounce these sounds differently from the standard. For example, one of my informants who speaks a nonstandard dialect of Armenian initially insisted to me that the palatalized consonants he was producing were *not* palatalized. The reason for this was that these particular palatalized sounds correspond to non-palatalized sounds in standard Armenian. Ironically, once I had demonstrated to his satisfaction that his sounds were in fact different from their standard Armenian equivalents (in this case, by pointing out the parallel to his other native language, Russian, which also contains both plain and palatalized consonants), he promptly began to scold me every time I had difficulty discerning whether a consonant he had just produced was palatalized or not!

To sum up, you cannot always count on your informants to produce reliable insights into the workings of their phonological systems. When this comes to pass, you are limited to making your own observations based on the lexicographical and other data you have collected. There are two keys to carrying out this procedure successfully: (1) look at individual sounds to see if they alternate in different phonological contexts or in different members of a paradigm; (2) look for common phonological processes, because some of these are bound to occur in your language. We consider the second strategy in more detail in section 5, and focus here on the first strategy.

When you are basing your analysis solely on data you have collected from your informants, you need to beware of productive phonological processes. Oftentimes, segments optionally undergo either a productive rule or a more restricted rule, and you want to avoid missing the latter type, which you are much less likely to hear employed by your informants. For example, the English word *hat* can be pronounced either as [*hæʔtʼ*] or as [*hætʰ*]; the former pronunciation results from the application of an optional rule of glottalization, and the latter results from the productive rule of aspiration that applies to stops when no other rules have applied. If you are having American informants produce words in isolation, you are likely to hear only [*hætʰ*], whereas if you are eliciting only casual connected speech, you are likely to hear only [*hæʔtʼ*], and never [*hætʰ*]. How can one go about finding rules that are restricted in their occurrence, and are often superseded by productive rules? As in similar situations discussed earlier, you can always try to read grammars of the language in advance, to see what sorts of rules are already known to exist. You can also ask your informants if they know of any other ways of producing a given form: asking this of an American a propos of [*hætʰ*] might elicit the alternate pronunciation [*hæʔtʼ*], for example. Rules that are more restricted in their application and distribution are also more likely to appear in connected speech than in careful speech, so it is a good idea to collect texts, conversations, and so on.

One last tip: resist the common temptation to fall into despair about finding alternations and rules in the language you are studying. One of the most common reactions in students examining a language for the first time is "this language is completely regular and systematic! It has no rules!" Your informants will generally be all too happy to contribute to this notion, for various reasons that will become clear to you [illegible]

3. How to Identify Phonological Constraints

Constraints can be a bit more difficult to identify than rules. The reason for this is that constraints are often static: they do not appear as alternations within a paradigm or changes to a segment in a specific environment, but rather they manifest themselves as generalizations over the entire vocabulary. In order to find constraints, you therefore need to identify possibilities that *could* exist, but do not. For example, students are taught in English-speaking schools that we form comparatives by adding *-er*, and superlatives by adding *-est*: *fast, faster, fastest*. What they are not taught is that there is a constraint which holds over these two morphemes. This constraint disallows the addition of *-er* and *-est* to words containing two or more syllables: **expensiver*, **expensivest*; **intelligenter*, **intelligentest*; **decayeder*, **decayedest*, **oftener*. (Some bisyllabic words that end in an unstressable syllable are also allowed, such as adjectives in *-y* and *-ow*: *happier*, *narrower*, etc.) English has many other constraints, including one that disallows verbs formed by adding *-ize* to adjectives with final stress (**corruptize*, etc.), one that disallows syllables beginning with *ŋ-*, one that disallows the vowels {ɪ ɛ ʊ} from occurring in word-final position, and so on.

If you are going to identify phonological constraints in the language you are studying, you will have to do a lot of work on your own time. It will help to pore over the data you have collected, looking for environments where a specific phoneme should be able to occur but doesn't, cases where you expect a certain morpheme to appear but it doesn't, and so on. We will not discuss these techniques in detail here, because by and large they do not require working with an informant, and the tricks involved are more properly the domain of an introductory textbook on phonology and morphology (we recommend Swadesh 1934, for example).

However, there are two constraint-related activities you can engage in with your informants. (1) You can directly ask them if they know of any constraints on the distribution of segments or morphemes in their language. This can be facilitated by giving examples of constraints in English or another language with which both you and they are familiar. The direct interrogation method normally does not work, but it can't hurt to try, and you may obtain some useful results. (2) You can use your informants to verify or disprove constraints that you think may exist in their language. This is an extremely rewarding technique, as it does not require informants to conduct any conscious analysis of their language. If you present your informants with a list of forms violating the constraint you have postulated, they will normally be able to say immediately whether these forms are acceptable or not. They may also be able to produce counterexamples immediately if you present them with a generalization you have come up with, such as "there are no words that begin with [ŋ]". We highly recommend this strategy, because it can save you dozens of hours looking for counterexamples yourself in dictionaries and grammars.

4. Constructing a Phonological Analysis

When you are constructing your phonological analysis by identifying rules and constraints, two major concerns will inevitably arise. The first involves the point at which you should begin constructing a phonological analysis relative to the point where you start collecting data. The second involves the intellectual and mechanical challenge of selecting underlying forms.

4.1. When to Conduct a Phonological Analysis

There are two schools of thought on the question of when to begin constructing a phonological analysis. One camp, of which the phonetician Peter Ladefoged is a notable

proponent, maintains that you should begin phonological analysis immediately when conducting field work (cf. Ladefoged 1997). In a certain sense, it is impossible to do otherwise, because one cannot transcribe without having conducted some amount of phonological analysis—one needs to have decided what sounds are different and what sounds are not, what to write down and what not to write down, how much detail to write down, and so on. Each of these decisions involves a fundamental sort of phonological analysis, without which field work would not be possible. Let us put aside this technical point, though, and assume instead that by phonological analysis we actually mean *conscious*, *intentional* phonological analysis. There are at least two advantages to beginning this sort of analysis as soon as you start collecting field data. (1) When you analyze your data immediately, it is much easier to know what questions to ask next, or at least in the next session. Explicit formulation of an analysis has a way of revealing flaws in your thought processes that would otherwise have remained blissfully and dangerously unnoticed. These flaws can then be tested and corrected through further questioning of your informants. (2) If you go over your data immediately, you can check your transcriptions for consistency. Especially in the first few sessions, you are likely to mishear many forms, unless you check all data rigorously with your informants as soon as you collect them. By looking over your field notes, you may be able to catch some of your transcription errors when you compare related forms. The problem with this strategy, though, is that you may go too far in your corrections, and remove variations that result from the application of phonological rules, and are not actually mistakes in transcription.

This problem is the *pou sto* of the other school of thought, which maintains that it is not a good idea to begin phonological analysis right away. It is better, the reasoning goes, to let the data take you where they will, unhindered by your biases and expectations. This strategy can work very well for field workers with excellent ears, but needs to be tempered somewhat for individuals who have difficulties with transcription. Linguists who can only work with a particular set of informants for a limited period of time should also steer clear of the late analysis philosophy, in order to avoid discovering after their time has elapsed that they missed or mistranscribed key information.

4.2. How to pick an underlying form

When you finally get down to the business of analyzing the data you have collected, one of the most fundamental questions you will have to grapple with is how to pick an underlying form for a given segment, morpheme, or word. (In case you are wondering why it is necessary to bother coming up with underlying forms when doing field work: having a phonemic analysis in your head that can generate appropriate underlying forms for the data that your informants produce can add consistency to your transcriptions, clarify what information you need to collect from your informants, help you create a phonemic writing system for the language, and so on.) Let us take a concrete example, the allophones of English /*t*/. Recall from our earlier discussion that there are five types of *t*-sounds: [*t*], [*tʰ*], [*ɾ*], [*ʔtʰ*], and [*ʧ*]. Assuming that these are all allophones of a single phoneme, how do we decide what the form of this phoneme is? In almost all cases, you want to select one of the surface allophones, rather than some other segment that never actually appears. With the five allophones above, for example, it does not make sense to postulate /*q*/ as the underlying phoneme from which they all are derived, unless you have extremely strong evidence in favor of such an abstract analysis. The particular allophone that you choose as the underlying form

changing /ʧ/ to [tʰ] at the beginning of stressed syllables, in order to account for words such as *tap* [tʰæp]. This rule is not only implausible phonetically, but also predicts that words such as *choose* should be pronounced *[tʰuz], and so on. It turns out in this particular case that the best candidate for the underlying form is /t/, because we can derive the other allophones from this via relatively straightforward rules. By "straightforward rules" we mean the type of rules that are common cross-linguistically, and plausible phonetically. We enumerate a number of rules of this type in section 5.

Beware that with some sets of allophones, it may not be easy to identify the underlying form. Turkish has a high vowel, for example, which surfaces as [i], [y], [ɨ], or [u], depending on the phonological context in which it occurs. Each of these allophones can be derived with equal ease from each of the other allophones, though, so it is extremely difficult to pick any one of them over the others as the underlying form. In cases like this, it is normally acceptable to shelve the task of picking a phonemic representation.

5. Common Phonological Processes

If you are having trouble identifying phonological rules and constraints in spite of the advice we have given so far, it can be very helpful to look out for common phonological phenomena. We provide below a list of phonological phenomena which are so common cross-linguistically that we guarantee you will find at least a few of them in whatever language you choose to study.

- Consonant-vowel interactions, such as **palatalization** (e.g. /k/ and /g/ become [c] and [ɟ] respectively before front vowels in English), **nasalization** (e.g. vowels become nasalized before nasal consonants in English), and **rounding** (e.g. English consonants are often pronounced with lip rounding before round vowels in the careful speech of older speakers).
- **Aspiration**, typically of voiceless consonants at the beginning of syllables, and sometimes of consonants at the end of words.
- **Unrelease** of consonants at the end of a syllable (i.e. before other consonants or at the end of a word).
- **Voicing** of consonants after nasals, between vowels, or adjacent to voiced consonants.
- **Devoicing** of consonants in word-initial position, in word-final position, or adjacent to voiceless consonants.
- **Simplification** of consonant clusters and affricates (e.g. the common American pronunciation of [*lɪfs*] for *lifts*, etc.).
- **Epenthesis** (insertion) of vowels in unpronounceable consonant clusters (e.g. English [*rɪðəm*] for *rhythm*), epenthesis of consonants either between two vowels, at the beginning of a vowel-initial word, or in certain kinds of consonant clusters (typically a nasal followed by a sonorant or a fricative, e.g. [*sʌmpθiŋ*] for *something*).
- **Glide insertion** between two vowels.
- **Deletion** of vowels next to other vowels, and of consonants next to other consonants or at the end of a word.
- **Assimilation** to the place of articulation of a following segment, particularly by nasals (cf. the *in-* morpheme discussed earlier).

We consider other common phenomena such as reduplication, stress shift, lengthening, and so on in the next chapter.