CHAPTER 2
How are phonological data created?

PREVIEW

This chapter explains how phonological data, especially transcriptions, are created. You will:

- experience how field linguists discover the transcriptional symbols to use for a language, exemplified in the language Logoori
- distinguish continuous variation in physical realization from grammatical changes in linguistic sound categories via rules of grammar

KEY TERMS

transcription

auditory judgment

variation

phonetic implementation

1. What Is Transcription?

The main point made in Chapter 1 is that phonology is about the analysis of language sound data, viewing sound as a discrete mental symbol. We reduce continuous physical sound in speech to a sequence of specialized symbols, the enumerable sounds of a language which can be printed in a book or texted on a phone – a transcription, which is the foundational data for phonological analysis. These transcriptional letters represent the cognitive symbols which we claim a (mental) grammar operates on. Throughout the book and in all phonological analysis, you will be dealing with sequences of such written symbols, language data. You should then wonder, *how* are these data created in the first place? On what basis are these data created – what thing in reality do they represent?

All of the data in this book, and any other writings on phonology, originate from somewhere else, such as a book or article, sometimes a personal data repository. They were originally created by someone, using an analytical procedure. In this book, some of what you will experience comes from my own personal data repository, the result of decades of study working with speakers of various languages of the world. Ultimately, phonological data come from a written source prepared by a linguist with personal experience in the language. The source may be something written by a native speaker scholar in the language, but the most enlightening case for us to consider – where the author has the least prior knowledge of the language – is what happens when a specialist in field linguistics, having no prior knowledge or preconceptions of the language, works with speakers of the language and learns enough about the language that they can provide us with data and analysis of the language, going from zero knowledge to publishable knowledge. The basis for all grammars (leaving aside grammars of long-dead languages which involve guesswork about pronunciation) is the analysis of speech, which requires reducing physical speech to fixed symbols.

You might think that reducing language to a form suitable for books is simple, we should just use the standard spelling of words in the language. This won't work for many reasons. First, very many languages of the world are not written, and there is no standard spelling or even widely-known script for the language. Second, as is well known from English (and many other languages), spelling conventions tend to be highly variable. The letter "c" means very different things in French, Chinese (Pinyin spelling), Somali, Zulu, Turkish, Lushootseed and Latvian. English spelling is particularly arbitrary, given the many ways of spelling the same vowel in "l**ee**k, l**ea**n, l**ie**n, L**e**na, am**oe**ba, sk**i**, rec**ei**ve". Languages in their written form very often leave out important aspects of pronunciation of interest to phonology, such as the difference between the noun "**i**mport" versus the verb "imp**o**rt" (stress on the first versus second syllable). Furthermore, it would be impractical for a student to have to learn dozens of unfamiliar scripts in order to understand and analyze data like ብቅዓት, مَفْعُوْل, हाथी, ղողալ or 송아지. For this reason, linguists present data using a compact system of symbols standing for the mental symbols that we believe underly language. Data is presented in the form of a transcription where each letter has a standardized auditory definition. This chapter walks through the logic of converting physical speech into written symbols, where we start with no preconceptions about the language in question.

You were probably introduced to the present standard system, the International Phonetic Alphabet, in an introductory course. Two-thirds of that system is the IPA chart <https://www.internationalphoneticassociation.org/sites/default/files/IPA2005_2000px.png inserted here>, which is a set of symbols and associated articulatory labels. The symbol [p] represents a voiceless bilabial plosive, which means that it is formed without vocal fold vibration, it has complete closure of both lips, and the lip closure is maintained throughout its production. It is not difficult for the researcher to verify these aspects of the production of [p] by observing the lips or feeling with your fingers for vibrations in the larynx, but without invasive physiological studies, it is difficult to directly observe the articulatory state of most language sounds, such as a voiceless epiglottal fricative (produced at the back of the throat). If we can't see what is going on when people speak, how can we have any idea whether to use [q], [ɢ], [ʀ], [ʁ], [ħ], [ʜ], [ʢ] or any of the other back-of-the-throat symbols, and how do we really know that it is a back-of-the-throat sound? Normally, we just ask the speakers who we are working with to pronounce words in their language, and we don't ask for instructions on how to articulate those sounds (speakers usually don't have a very good idea how they do what they do). How do children learn how to articulate the sounds of their language, since they don't have access to x-ray or ultrasound machinery?

We do not usually decide on transcriptional symbols by invasive physiological inspection. How *do* we do it? A third factor, the auditory value of the transcriptional symbol, provides the necessary linkage that allows us to relate a particular language sound to an appropriate IPA symbol and its articulatory description. To put it simply, each IPA symbol has a distinct characteristic sound. The vowel [ɑ] sounds different from the vowels [ɛ] or [æ], and the consonant [n] sounds different from [ŋ]. One of the basic tools and necessary trainings of the field phonologist is what is conventionally known as "ear training", a course of study where one gains knowledge of this standardized sound-to-symbol relationship, and develops the skill required to assign an appropriate phonetic symbol to a particular language sound.

Historically, transmission of knowledge of those sound-to-symbol standards was a bit haphazard, often reducing to crosslinguistic comparison. One might learn that [u] is "the vowel in English *s**u**per* or German *Z**u**g, Kr**u**g*", which is only helpful if you have experience with the relevant dialect of English or German, the one that I might be referring to. The actual pronunciation of "super" in English varies quite substantially, ranging from [supɹ̩], [sɪwpɹ̩], [supə], [sʉpə], and many other possibilities, so even if you know the word "super" in English, you don't necessarily know *which* pronunciation the author has in mind in saying "like super". Much more useful was being trained by a expert teacher who previously received standard training from another previously-trained expert teacher, and who could accurately produce all of those sounds. Contemporarily, a simple solution is that you can consult online collections of reference pronunciations produced by experts in the IPA, found at <https://www.internationalphoneticassociation.org/IPAcharts/inter_chart_2018/IPA_2018.html>, a resource which we will rely on in this chapter. The four scholars providing reference pronunciations there have extensive experience in the IPA, and here we will treat these recordings as defining auditory standards for IPA symbols (it is important to understand that the IPA does not have official pronunciations, but the association has set these forth as exemplars).

This chapter simply seeks to give you basic awareness of how a field phonologist converts continuous speech into a set of discrete symbols, and does not aspire to teach all of the requisite skills for doing this yourself, for all sounds in any language. Those are skills developed in field methods courses and in the course of years of elicitation experience with particular languages. We will consider some basic problems of symbol choice for vowels, because vowels are very easy in some ways (they can be prolonged and uttered alone, which aids hearing) and very difficult in other ways (they are highly variable). This chapter works in parallel with online materials <interactive page https://languagedescriptions.github.io/IP3/Ch2.html. Zip file archive to be placed on Zenodo when this becomes finalized>.

Before looking at our first case study from Logoori, we start basic warm-up practice in judging vowels. For this, you should listen to the vowel recordings on the IPA web page at https://www.internationalphoneticassociation.org/IPAcharts/inter_chart_2018/IPA_2018.html. Calibrate your vowel experience by listening to the recordings made by expert phoneticians, at least for the vowels [i ɪ e ɛ æ a], so that you have experience with the standard pronunciations of these vowel symbols. Bookmark the table so that you can easily call up the reference recordings. The question to consider in listening to these recordings is, are the online reference pronunciations identical within a given symbol, or can you hear any differences?

The easy answer is that all of the recordings sound somewhat different from each other. At the same time, within the specific symbolic group, each of the recordings sounds similar to the other 3 recordings within the group, and each group collectively sounds different from the examples in other symbolic groups. I find all 4 recordings of [i] to be very similar, although I hear differences in personal voice quality, and there are also uncontrolled differences in intonation (duration and pitch) which are not part of the defining nature of the [i] auditory samples. Pitch is clearly falling in the samples of [i] for JE and JH compared to PL and JW, but we just ignore that difference as non-essential variation.

Likewise the samples of [ɪ] are highly similar within that group. The small variation within each symbolic group is not significant for the purposes of the field phonologist, what is important is the significant difference *between* the symbolic groups. As a whole, the [i] examples sound different compared to the [ɪ] examples, and the [e] examples as a group sound different from the [ɛ] examples as a group. Transcription is, at its core, the enterprise of detecting similarities in some respect, and setting aside differences in other respects. Of course, the ease with which you might hear differences also depends on the languages that you speak. A person speaking Quechua, Tamazight or Inuit (which have 3 vowels) will have a reduced native-language basis for distinguishing the 9 vowels of Sotho, compared to a speaker of English, which has a more-similar vowel system.

What underlies these differences in vowels is a physical difference (the primary one for vowels), namely the formant values. The first and second formants of these reference recordings of [i] are rather far apart, compared to those for [ɪ] where the formants are closer together. The placement of a particular vowel in a typical vowel chart reflects these formant relations. When we say that one vowel sounds "higher" than another, this generally refers to the value of the first formant (F_1) – higher vowels have lower F_1, and the impression of differences in backness / frontness or rounding generally reflects differences of the second formant (F_2). We interpolate from an auditory standard to an articulatory classification, based on prior research into the production of particular language sounds.

At this point, we are only considering auditory comparisons between a specific instance of a vowel in a specific language, in relation to a fixed standard. Our goal is to make a choice from the set of phonetic symbols, picking out the symbol that most-closely matches the defined standard, given the specific pronunciation. The problem that immediately arises in comparing an utterance in a language to the standard pronunciation samples is that specific language samples are typically outside the narrow range of variation encountered in those reference pronunciations. A typical American English [i] is not as high- and front-sounding as the reference samples for IPA [i], but American English [i] is still closer to reference sample [i] than it is to any other vowel. The Turkish vowel transcribed in IPA as [i] is physically midway between the Dutch vowel transcribed as [i] and the Dutch vowel transcribed as [ɪ]. Thus speakers of Turkish are aiming for a different physical sound target when they produce [i] in Turkish, compared to what Dutch speakers are aiming for when they produce [i] or [ɪ] (see Ahn & Chodroff data: https://osf.io/t957v/files/osfstorage/656b65cd783ec60ebaeddc37). In selecting a vowel symbol, we aim to find the *closest* match. A vowel-category measurement in the form of an IPA symbol choice is not a numeric formant measurement, which we touched on in the first chapter.

To summarize, when a field phonologist begins the process of creating a collection of written data for a language, they start by comparing specific sounds of the language to fixed reference values. Those reference values are broad ranges, not exact points, and the fundamental question for the transcriber is "is the vowel in this utterance closer to standard sample [i], or is it closer to [ɪ]?", or some other applicable choice. Our next task is to actually do this, by discovering the vowels of Logoori <https://languagedescriptions.github.io/IP3/Ch2.html – please note that these URLs after "github.io" are case-sensitive>. For this section, you need to follow along with the web page.

2. Logoori Vowels

Logoori is a member of the Bantu language family, spoken primarily in western Kenya, a language which I have been researching for a number of years. This section partially re-creates my process of learning to transcribe vowel differences, specifically drawing from the speech of one speaker, Franklin Inzuga. The usual procedure for gathering such data is via interviews, where I would ask "How do you say ___ in Logoori?", filling in various English words, then recording the answer. By "recording", I mean literally making a sound recording of what he says, so that I can play the response repeatedly, and consider pronunciation carefully given this permanent record. In this instance, the procedure was a bit different from my usual since the speaker lives in Kenya and I live in the US, and we communicated indirectly via the internet – I provide a written list of questions, he makes recordings, which I then transcribe by listening to what he said and comparing that to the defined auditory standard. Because of my prior experience with the language, I was able to go directly to more-informative words.

One does not typically start research on a language by asking for a word meaning "to stop a ball", "to pick crops" or "to use a quail as bait", even though those are words exemplified here. What we will see here is a simplified, condensed and curated version of the field-phonologists's experience, recreating the logic of converting auditory experiences to phonological transcriptions, mainly focusing on one question: what are the vowels of Logoori? All of the words below are online for you to listen to, and you should follow along and make your own judgments.

In this section, we focus just on vowel identification, ignoring various interesting facts about consonants. In actual practice, the fieldworker has to juggle all aspects of transcription at the same time. In this chapter, we delay certain problems of consonants to a later section when we face the topic of variability.

2.1. Set 1 recordings

Our first collection of Logoori words is the Set 1 recordings online, which is an initial collection of words of the language. Listen to those words, and compare what you hear to the suggested initial transcriptions. This initial hypothesis about the words (the most likely conclusion that you would reach based just on listening to this list) is given in (1). The consonants of Logoori are not particularly difficult, but so as to not ignore *all* interesting consonant problems, you will notice that these Logoori examples include the retroflex lateral consonant [ɭ] in words like [kutaɭa, kuɭuma]. This consonant sounds like a mix of *l* and *r*. Usually, the linguist would puzzle over this sound because it is rare in the languages of the world, though it may be ordinary to native speakers of Korean, Tamil or certain dialects of Norwegian. To simplify matters I am telling you that that is [ɭ], which you would probably eventually realize yourself by comparing these Logoori words with the IPA reference recordings for various laterals and rhotics.

Since Logoori is a Bantu language and it is known that Bantu languages overwhelmingly have either 5 or 7 vowel qualities ([a], and 2 or 3 each of front and back vowels), we have an expectation of the vowels that we will find. This is an expectation, not a rule: your conclusions must be based on what you directly experience. When listening to these samples, remember that your reference point is not English, Chinese or any other specific language, it is the expert pronunciations in the online IPA table which are designed to guide analysis of all languages.

(1) *Set 1 recording*

kudasa	to increase
kudaja	to demand payment
kutaɭa	to pass rumors
kudamba	to be vulnerable
kɔgɛnja	to wonder
kɔmɛda	to add to
kɔdɛɾa	to stop a ball
kɔvɛga	to shave
kusinja	to annoy
kusiɾa	to jump
kuvita	to pass
kuɭima	to plow
kugika	to exaggerate
kuhiɾa	to drive
kuvisa	to hide something
kɔɭɔta	to dream
kɔsɔɭa	to protect
kɔkɔɭa	to do
kɔɭɔmba	to make
kuguta	to stack up
kukunja	to be stunted
kutuɭa	to smelt
kuɭuma	to bite
kusuɭa	to refuse
kuguta	to win

Theses examples are all verb infinitives, which have a prefix with *k* plus a back round vowel *u* or *ɔ*, and they all end in the vowel *a*. This tight control over the data allows us to focus on one question: what is the second vowel of the word? With a large enough sample of such words, we will get data relevant to answering a fundamental phonological question "What are the vowels of Logoori?". The data in (1) are already sorted in accordance with the very question that we seek to answer, although in actual practice examples tend to be collected haphazardly, until one understands the language well enough to collect data in an orderly fashion.

We now have an initial hypothesis about vowels and the words which have them. If the transcriptions of (1) are correct, each example of *ɛ* should sound like every other example of *ɛ*, each example of *i* should sound like every other example of *i* and so on. A virtue of recording examples and arranging them into a playable table as we find in the online data is that you can selectively compare words, asking "does the second vowel of 'to add' sound the same as the second vowel of 'to stop a ball'?", likewise comparing 'to be vulnerable' and 'to increase' for sameness of vowels. If *kudamba, kudaja, kudasa* and *kutaɭa* likewise *kɔmɛda, kɔdɛra, kɔgɛnja* and *kɔvɛga* are the same in having *a* and *ɛ* as their respective second vowels, the vowels in each of those two subgroups should sound the same, an idea that we test by pairwise listening of the suspected [ɛ] words of the [a] words. You should now do this comparative listening on the web page.

Although the basic division of vowel qualities that you probably arrived at for these words as in (1) seems correct, after listening a second or third time to these words, especially comparing examples of "a", "i", "ɔ" and so on for similarity, you will most likely detect two other perceptible phonetic factors which were *not* initially obvious, and were not noted in the first transcriptions in (1), properties which further subdivide the vowels. We notice that the second vowels in *kudamba, kudaja* are much longer than those of *kudasa, kutaɭa*, likewise the vowels of *kɔmɛda, kɔdɛra* are longer than those of *kɔgɛnja, kɔvɛga*. The importance of vowel length becomes especially clear when you compare the two words originally transcribed [kuguta]: "to win" and "to stack up" which are very similar in sound, yet they are not the same. If you want, you can even measure the difference of duration in those vowels, however here we are concerned with auditory judgments, which is the information that is also available to a child learning the language. We therefore consider the alternative possibility that the language has not just *ɛ, a* etc. but also *ɛ:, a:*. We would surmise that the words actually are *kuda:mba, kuda:ja, kudasa* and *kutaɭa*, likewise *kɔmɛ:da, kɔdɛ:ra, kɔgɛnja, kɔvɛga*. Vowel length seems to be an orthogonal pronunciation property of vowels, independent of basic quality. Now being aware of this possibility, we should refine all of our transcriptions.

In the course of this review (by re-listening to the examples in Set 1 based on this new understanding), we also notice a second dimension of vowel sub-classification which may not have been initially noticed: voice pitch can be different between words. Pitch is higher in *kuda:mba* and *kutaɭa* than it is in *kudasa* and *kuda:ja*. This difference is also found in *kɔvɛga* where the second vowel has a higher pitch than in *kɔmɛ:da, kɔdɛ:ra, kɔgɛnja*. At this earliest stage of analyzing the language, we do not decide if this is to be treated as "tone", "stress" or "accent" because that is a question of advanced analysis, and we are focused just on accurate fact-reporting. We simply note that there is some second difference in vowels, and focus auditorily on pitch: higher versus lower. We will use an accent mark on the vowel to indicate this difference, e.g. *kudá:mba, kutáɭa*, *kuda:ja*, *kudasa*. In light of these hypotheses, we arrive at a revised set of transcriptions for the recordings, with a revised pairing of transcription and recording online in Set 1 Corrections, where you can directly match the improved transcriptions with these same recordings, reinforcing the auditory import of length and tone marks.

(2) *Set 1 Corrections: Tone and Vowel Length*

kudasa	to increase
kuda:ja	to demand payment
kútáɭa	to pass rumors
kúdá:mba	to be vulnerable
kɔgɛnja	to wonder
kɔmɛ:da	to add to
kɔdɛ:ɾa	to stop a ball
kɔ́vɛ́ga	to shave
kusinja	to annoy
kusi:ɾa	to jump
kúvíta	to pass
kuɭima	to plow
kugí:ka	to exaggerate
kúhíɾa	to drive
kúvísa	to hide something
kɔkɔja	to gather goats or sheep together
kɔsɔ:ɭa	to protect
kɔkɔ́ɭa	to do
kɔ́ɭɔ́:mba	to make
kugúta	to stack up
kuku:nja	to be stunted
kútúɭa	to smelt
kúɭúma	to bite
kúsú:ɭa	to refuse
kugú:ta	to win

Transcription is thus the art of auditory subclassification arrived at by repeated listening to diverse speech samples, paying attention to differences and similarities. This art relies heavily on review and hypothesis-revision.

The student should be aware that transcribing tone is often more difficult than transcribing segmental qualities, because tone is a relative property of having "higher" or "lower" pitch – compared to what? The range of variation in distinguishing vowel qualities is much more limited. Therefore, judgments that a vowel has "higher" versus "lower" pitch are much harder to make, and you are likely to be more uncertain about tonal transcriptions compared to segment-quality transcriptions, based on the limited data available here.

Based on these data, we have observed the following vowels in Logoori.

(3)

a	a:	á	á:
ɛ	ɛ:	ɛ́	
i	i:	í	í:
ɔ	ɔ:	ɔ́	ɔ́:
u	u:	ú	ú:

For each known vowel quality of the language (and because we were lucky in collecting a sufficiently-diverse set of words), we have founnd four subtypes: long versus short, and high-pitched versus low-pitched, except that we have no example of [ɛ́:]. This may be a systematic fact of the language if [ɛ́:] does not exist, or it may be an accident of the specific collection of words.

2.2. Set 2 recordings

We further test our hypothesis about vowels by increasing our supply of recordings and checking whether new examples present any additional vowel types – now listen to the online Set 2 recordings. An important part of such re-checking is to gather second pronunciations of some already-collected words, which should not result in a different vowel. As a step towards re-checking our earlier conclusion, we include second pronunciations of a few earlier words such as [kudasa, kɔkɔ́ɭa, kuɭima] and [kɔdɛ:ɾa].

(4) *Set 2 Recordings*

kudasa	to increase
kɔ́kɔ́ɭa	to do
kuɭima	to plow
kɔdɛ:ɾa	to stop a ball
kuva:dza	to carve
kuhama	to move residence
kɔ́hɛ́:nda	to worry
kuhi:nda	to protrude
kuɭi:ma	to use a quail as bait
kusi:ŋga	to give a bath
kuhi:nda	to give excessive amounts
kúsí:nza	to slaughter
kɔnɔga	to pluck
kuɭuma	to make noise
kúdú:mba	to push
kútúɭa	to come out
kuɭuha	to be tired
kuɭu:nda	to gather in a crowded space
kúkú:nja	to walk majestically

On the web page under Word-repetitions, the two repetitions of [kʊdasa, kɔ́kɔ́ɭa, kuɭima, kɔdɛ:ɾa] are arranged side-by side for easy listening: listen to these pairs to hear the similarities and the differences. Although there are some subtle physical differences in pronunciation between these repetitions (speech rate in the two examples of [kudasa], an overall difference in [kuɭima] resulting from the fact that the two examples were recorded in a different room), one particular difference stands out as potentially significant. Of the two pronunciations of ‘to do’, in the first pronunciation the vowel of the initial syllable has a lower pitch compared to the second vowel, but in the second pronunciation the pitch of the first two syllables sounds as high as that of the second syllable, a fact that we note in our transcription by transcribing the first pronunciation as [kɔkɔ́ɭa] and the second as [kɔ́kɔ́ɭa]. This introduces a possibly phonological question: *why* do the two pronunciations sound so different? Because our focus is accurate recording as the foundation for analysis, we do not attempt to answer the analytic question right now, instead we just note the fact.

Listening carefully to these new words, we also detect a previously un-noticed difference in vowels, going beyond tone and vowel length. While we originally transcribed [i] as the second vowel in [kuhi:nda, kuɭi:ma, kusi:ŋga, kuhi:nda, kúsí:nza], the vowels in these examples do *not* all sound the same. The vowels of the first three, [kuhi:nda, kuɭi:ma, kusi:ŋga] sound very similar and like English "seen", but the vowels of [kuhi:nda, kúsí:nza] where [kuhi:nda] means 'to give excessive amounts' sound different, more like English "sin". In fact, it is a puzzle of sorts that we have two words with different meanings and different yet similar pronunciations – [kuhi:nda] 'to give excessive amounts' versus [kuhi:nda] 'to protrude'. FI confirms that these words *are* pronounced differently, and upon careful listening and comparison we realize that 'to give excessive amounts' should more be accurately transcribed as [kuhɪ:nda], but 'to protrude' is [kuhi:nda]. We likewise detect a difference within the vowels earlier transcribed as [u], that the vowel of [kuɭuma] 'to make a roaring noise' is [u] but that of [kuɭʊha] 'to be tired' is [ʊ]. As well, the vowel of [kuɭuma] 'to make a roaring sound' differs from [kʊ́ɭʊ́ma] 'to bite' in the first set both in tone and vowel quality.

Field linguists often use medium-technology computer methods to help them focus on challenging sound differences. In listening to a word like [kusi:ɾa] 'to jump' and a word like [kʊgí:ka] 'to exaggerate', the currently-relevant question is whether the second vowels in the two words are the same or different. The difficulty in making that judgment is that many other orthogonal differences in the words also catch our attention. A simple technological solution is to extract just the vowel part of these two words, and compare those vowel parts. Vowel portions of these words are therefore available online to listen to. As you can hear, the vowels extracted from [kusi:ɾa] and [kʊgí:ka] sound different. To be sure that we are not over-focused on a trivial physical difference in specific tokens, we also extract the vowels of [kuhi:nda] 'to protrude' and [kʊ́sí:nza] 'to slaughter', where we can focus on the sameness of the vowels [ɪ] versus [i].

Now that we are aware of the quality difference between [i, u] versus [ɪ, ʊ], we re-consider our prior transcriptions of [i, u]. By comparative listening within the words thought to have [u] and those thought to have [i], and being aware that [kuhɪ:nda] 'to give excessive amounts' and [kuhi:nda] 'to protrude', also [kuɭʊha] 'to be tired' and [kuɭuma] 'to make a roaring noise' have different vowels, we can re-evaluate the transcriptions of what we formerly thought were just the two vowels [i, u]. While re-checking the second-syllable vowel of these verbs, we also notice a difference in the pronunciation of the first syllable, the infinitive prefix. In some cases the prefix vowel is [u] and in other cases it is [ʊ] (we already transcribed but were not worried about explaining differences in the prefix vowels in [kɔ́kɔ́ɭa] and [kʊdasa]). This review of previous cases of *i, u* results in amended transcriptions, with corrected vowels matched with recordings online.

(5) *Vowel Quality Corrections: i vs. ɪ, u vs. ʊ*

kʊda:ja	to demand payment
kʊ́dá:mba	to be vulnerable
kʊdasa	to increase
kʊhama	to move residence
kʊ́táɭa	to pass rumors
kʊva:dza	to carve
kɔdɛ:ɾa	to stop a ball
kɔgɛnja	to wonder
kɔ́hɛ́:nda	to worry
kɔmɛ:da	to increase by adding
kɔ́vɛ́ga	to shave
kuhi:nda	to protrude
kúhíɾa	to drive
kuɭi:ma	to use a quail as bait
kusi:ŋga	to give a bath
kusi:ɾa	to jump
kusinja	to annoy
kúvísa	to hide something
kʊgí:ka	to exaggerate
kʊhɪ:nda	to give excessive amounts
kʊɭɪma	to plow
kʊ́sí:nza	to slaughter
kʊ́víta	to pass
kɔkɔ́ɭa, kɔ́kɔ́ɭa	to do
kɔ́ɭɔ́:mba	to make
kɔɭɔ́:ta	to dream
kɔnɔga	to pluck
kɔsɔ:ɭa	to protect
kúdú:mba	to push
kugú:ta	to win
kuku:nja	to be stunted
kuɭuma	to make noise
kútúɭa	to smelt
kʊgʊ́ta	to stack up
kʊ́kʊ́:nja	to walk majestically
kʊɭʊ:nda	to gather in a crowded space
kʊɭʊha	to be tired
kʊ́ɭʊ́ma	to bite
kʊ́sʊ́:ɭa	to refuse
kʊ́tʊ́ɭa	to come out

From this collection of 44 recordings, we now have a good initial idea of what the vowels of Logoori are. In terms of qualities, we have found seven types [i, ɪ, ɛ, a, ɔ, ʊ, u], as well as a difference between long and short vowels, and high-pitched and low-pitched vowels.

2.3. Exercise 1

In further investigating the language, we will need to expand the database as much as possible, to be more confident that we have uncovered all of the vowels. For present purposes, we expand the database just a bit by adding 16 more examples which are similar to words that we already have. Transcriptions are provided online for the first 8 of these words which will reinforce your perception of these vowel differences. As an exercise, you should provide transcriptions for the second 8 words in this collection.

2.4. Choice of symbol

The vowel judgments embodied in these transcriptions reflect the fact that, through comparative listening to the language samples over time and to the reference recordings, I find that the vowels which I transcribe as [i, u] are to me most like IPA standard [i, u]. There are also similar-sounding yet different vowels which I have written as [ɪ, ʊ]. The primary concern has been detecting those vowel sub-groups which are clearly distinct, as in [kuvísa] versus [kuvíta] or [kuku:nja] and [kukʊ́:nja]. These vowels are unquestionably different things in the language (which is not to say that detecting the difference is trivial, and the present introduction may be too abbreviated for a student to confidently perceive these distinctions). The secondary concern is most-accurately matching those choices to a fixed palette of possibilities – is a particular vowel more like IPA-ideal [i], [ɪ], [e] or [ɛ]? There is flexibility in symbol choice, so we could instead write [kuvísa, kuvéta, kuku:nja] and [kukó:nja]. There is also flexibility in how we might transcribe the vowels of [kɔmɛ:da] and [kɔsɔ:ɽa], we might instead have written [kome:da] and [koso:ɽa]. Appeal to the IPA reference pronunciations is not particularly helpful here, because the vowels as encountered in Logoori reside somewhere between reference [e o] and [ɛ ɔ]. We can tell that the front and back vowels come in three different heights, a highest, a middle and lowest height, but we have a hard time at this point saying *which* specific reference vowels the Logoori examples are closest to.

To be explicit, there are at least 4 imaginable analyses of these data.

(6)

A	B	C	*D	
kuvísa	kuvísa	kuvísa	kuvísa	‘to hide’
kuvéta	kuvíta	kuvíta	kuvéta	‘to pass’
kɔ́vɛ́ga	kɔ́vɛ́ga	kóvéga	kɔ́vɛ́ga	‘to shave’
kudú:mba	kudú:mba	kudú:mba	kudʊ́:mba	‘to push’
kukó:nja	kukʊ́:nja	kukʊ́:nja	kukó:nja	‘to walk majestically’
kɔ́ɽɔ́:mba	kɔ́ɽɔ́:mba	kóɽó:mba	kɔ́ɽɔ́:mba	‘to make’

We reject *D which posits [kuvísa] and [kudʊ́:mba], because auditorily the vowels of these words are a very good match to IPA [i, u] and a poor match to [ɪ, ʊ]. The other analyses remain plausible because e.g. the vowel of “to pass” is reasonably close to both [ɪ] and [e] (but not [i] or [ɛ]), and the vowel of “to shave” is reasonably close to both [e] and [ɛ] (but not [i] or [ɪ]). We might not be able to decide between A, B or C based solely on the physical sounds – but we will see decisive evidence below.

The field phonologist is unsatisfied with a small, excessively-controlled sample of words as the only basis for making broad claims about properties of the language. In carefully controlling the examples, we may have excluded some important type of word, so the next step is to expand the class of words under investigation a little, this time changing the vowel at the end of the word, picking a word form which has *i* as the last vowel rather than *a* – this gives us the words in Set 3 online. We arrived at this possibility for making words by accident, simply by asking curiosity questions which might reveal a new kind of grammatical structure.

(7) *Set 3 Recordings*

kʊnáví	we sewed
kʊdá:ji	we demanded payment
kʊda:mbi	we were vulnerable
kʊdásí	we increased
komé:di	we increased
kogé:ndi	we walked
kohe:ndi	we worried
kógéɾí	we measured
kʊtʃíví	we mixed
kʊsí:ndʒi	we gave a bath
kʊví:di	we sprinkled on medicine
kʊhí:ndi	we protruded
kʊhɪ:mi	we hunted
kʊsɪdʒi	we showed off
kʊsɪ:nzi	we slaughtered
kʊɭímí	we plowed
kʊhí:ndi	we gave excessive amounts
koɭo:mbi	we made
koɭo:ti	we dreamed
kogódí	we bent
kogó:si	we were polite
kuɭú:mbi	we pushed
kʊbú:dʒi	we tossed in bed
kuɭudʒi	we made ugali
kʊtuɭi	we smelted
kʊɭʊ́:ndi	we gathered in a crowded space
kʊvʊɭi	we lacked
kʊtʊɭi	we came out
kʊvʊ:tʃi	we woke up

After listening to these new examples and comparing them to previous ones, we observe a distinction in vowels which we had not seen in the earlier data. Examples of these new vowels, [e, o], are encountered in [koɭo:mbi] and [kohe:ndi]. The new vowels are similar to both [ɛ] and [ɪ], [ɔ] and [ʊ], yet are still different from those vowels. Previously, we were unsure whether the second-highest vowels are closer to [ɪ, ʊ] or [e, o], and whether the lowest vowels are [e, o] or [ɛ, ɔ]. With this new data, have an answer, because there are *four* pronounced vowel heights, not three. We have the following (listen online to the vowel comparisons under Set 3).

(8) *Vowel Comparisons*

kuhi:nda	to protrude
kʊhɪ:nda	to give excess
kohe:ndi	we worried
kɔgɛ:nda	to walk
kuku:nja	to be stunted
kʊ́kʊ́:nja	to walk majestically
koɭo̞:mbi	we made
kɔɭɔ́:ta	to dream

Although we were primarily concerned with basic vowel quality here, accuracy of tone and vowel length is also important to the description of Logoori. In the course of comparing the most recent set of words versus earlier words, you should have noticed a tonal difference in words with a H tone on a long syllable. There is a clear difference between the penultimate tone of the words in (9a) versus (9b), where pitch falls at the end of the second syllable in (a), but not in (b). These words are set out online under Tone Corrections for you to listen to.

(9) a. *Falling pitch*

kɔ́hɛ̂:nda	to worry
kɔkɔ̂:nja	to help
kʊ́kʊ́:nja	to walk majestically
kugû:ta	to win

b. *Level pitch*

komé:di	we increased
kʊdá:ji	we demanded payment
kʊsí:ndʒi	we gave a bath
kʊɭʊ́:ndi	we gathered in a crowded space

This new knowledge motivates expansion of our trancription and recognition of another distinction, where we distinguish high level versus falling pitch on vowels. We notate that property with an acute accent for level pitch and a circumflex accent for falling pitch. This results in another review of the data, now that we are aware of its existence, and yields our final re-transcription of words. Transcriptions again have to be augmented to make this distinction – the affected words are in (10).

(10) *Tone Corrections*

kɔ́hɛ̂:nda	to worry
kɔkɔ̂:nja	to help
kɔ́ɭɔ̂:mba	to make
kɔɭɔ̂:ta	to dream
kʊ́dâ:mba	to be vulnerable

kúdû:mba	to push
kʊgî:ka	to exaggerate
kugû:ta	to win
kʊ́kɔ̂:nja	to walk majestically
kʊ́sî:nza	to slaughter
kʊ́sɔ̂:ɭa	to refuse

It is of some interest that all examples of falling raised pitch occur in infinitives and all examples of level pitch on long vowels are in past tense verbs, but also exist in the infinitive on short vowels. Understanding why that might be so is a question of analysis, which calls for more data. What is immediately important is to *observe* that there exists such a difference in tone-types, and to know that in recording data, we must watch for that difference, just as we watch for the differences [i, ɪ] or [e, ɛ].

2.5. Exercise 2

As a final exercise in transcription, you should provide transcriptions of the words in online Exercise 2.

2.6. Vowel summary

We have now uncovered all of the phonological vowel qualities of Logoori, having discovered [i ɪ e ɛ u ʊ o ɔ a]. We have also found two orthogonal properties that combine with these vowel properties: vowel length and tone. For every vowel quality there may be long and short vowels, and vowels can have a higher tone, or not, furthermore higher tone comes in two types (level and falling). Although that rich system of 36 vowel types may seem like a heavy burden for the field-worker to shoulder, it is actually very simple if you focus on individual phonetic properties: particular tone, vowel length and the separate aspects of individual vowel quality such as front / back, high / mid / low, and tenseness.

Of course the linguist describing the language does not intuitively *know* that these are all of the vowels, until they have worked more extensively on the language. One can in practical linguistics only firmly reach such a conclusion after inspecting substantially more words. The vowel set [i ɪ e ɛ u ʊ o ɔ a] is a scientific hypothesis for Logoori, and a sensible one given expectations of symmetry, which is being constantly tested by adding more data to the repository, to see if any falsifying data occurs.

> The hypothesized set [i ɪ e ɛ u ʊ o ɔ a] is the result of falsifying an earlier hypothesis that the vowels were just [i ɪ ɛ u ʊ ɔ a], an understanding which derived from attending to words like [komé:di] alongside [kɔmɛ:da]

Having worked on the language for over a dozen years I am now reasonably confident that I will not uncover a new vowel like [ø] or [æ], and I take the hypothesis [i ɪ e ɛ u ʊ o ɔ a] to be an established fact. Phonological analyses such as you will encounter in this book is founded on the assumption that authors *have* accurately recorded and adequately investigated the facts.

There are some issues regarding vowels which we have not addressed, but they are "why" question of analysis, not "what" questions of data recording. For example, *why* do we have both [kɔkɔ́ɭa] and [kɔ́kɔ́ɭa] as transcriptions of 'to do' (depending on the specific recording). Why is the infinitive prefix [kɔ] sometimes and [kʊ] other times (also [ku])? Answering those questions moves us from transcription to phonological analysis.

3. Logoori Consonants and Variation

Various questions could have been raised about consonants, some of which which we now consider. We already noted that there is a consonant [ɭ] encountered in [kuɭi:ma] 'to use a quail as bait', [kʊɭɪma] 'to plow', [kútúɭa] 'to smelt', [kʊ́ɭʊ́ma] 'to bite', [kʊ́sʊ̂:ɭa] 'to refuse'. It is different from the flap *ɾ* which we encountered in [kɔdɛ:ɾa] 'to stop a ball', [kusi:ɾa] 'to jump' or [kúhíɾa] 'to drive'.

Another challenging consonant is [g] in [kɔ**g**ɛnja] 'to wonder', [kɔ**g**ɔ:nja] 'to persuade', [kʊ**g**í:mbi] 'we made rain', [kʊ**g**ɪ:t^{ʃ}i] 'we exaggerated'. Sometimes *g* sounds like ordinary [g], sometimes it sounds like the voiced velar fricative [ɣ] perhaps familiar from Spanish. The same issue arises with *k* in many examples above such as [**k**ʊgíní] 'we falsely accused', [**k**ugû:ta] 'to win' or [**k**ɔ**k**ɔja] 'to gather goats or sheep together'. Sometimes that consonant sounds like [k], sometimes it sounds somewhat like [x]. Perhaps we should look closer at the measurable physical reality, not just depending on auditory response to listening to examples. The question which we hope to answer is whether these consonants are physically realized as stops (plosives), or as fricatives. The deeper reason for undertaking an acoustic investigation of these consonants is to concretize different notions of variation, a concept which underlies linguistic analysis.

3.1. The phonetics of stops and fricatives

The physical difference between stops and fricatives is as follows. Ideally in a stop, a complete obstruction of airflow through the vocal tract is created and maintained for some period of time. This allows build-up of air pressure, then the obstruction is released. Upon release of [g], we expect a rapid increase of airflow rushing past the point of obstruction. In a fricative like [ɣ], there is a narrowing of the vocal tract but not a complete obstruction, so there is much less increase of air pressure behind that narrowing. Forming and release this narrowing is slower than the complete on / off closure and opening of the vocal tract made during the production of a stop. There is even a third phonetically-recognized degree of construction, an "approximant" where the velar approximant, a sound with very little obstruction of airflow, is written as [ɰ]. A child learning a language has to figure out how to articulate the sounds of a language from what they hear, because they have no access to x-ray or ultra-sound technology. They also do not have computers for acoustic analysis or professional training in getting from acoustic outputs to reconstructed articulatory cause. We don't know how children do what they do, we just know that somehow they figure out how to sound like others around them. We too can use acoustic facts as a basis for understanding articulation.

Consonants standing between vowels best-exemplify the essential acoustic properties of different consonant constrictions. In the displays which we see below, the acoustic waveform is on top and there is a spectrogram below it. The waveform is a continuous graph of variations in sound pressure over time, where the vertical excursions from the center line may be large or small at a given time. In the spectrogram below the waveform, we see the relative intensity (represented via darkness) at different frequencies (the vertical axis) over time (the horizontal axis). Above the waveform we see a kind of intensity display, one designed to reflect rapid changes in acoustic energy, a display which hugs the outline of the waveform (the following pictures are paired with the recorded sound online under Consonant Variation). At the very bottom is the segmentation, which is a judgment of where the individual sounds as discrete units occur. Sometimes there are clear visual landmarks pointing to segment boundaries, and sometimes the exact boundaries are very unclear. We often rely on listening to sub-parts of the recording in order to resolve conflicting criteria. In the examples below, my practice is to include as much information in the signal as can be attributed to the consonant in question as being part of the consonant.

We start with [komé:di] 'we increased' and [kʊdasa] 'to increase', illustrating [d] between vowels (graphs and recordings are online under Constriction Basics).

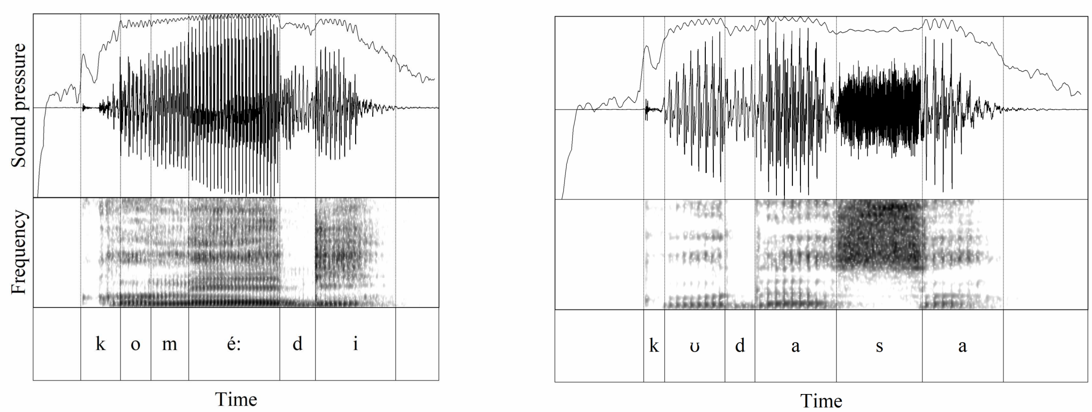

Figure 1

During the portions labeled as vowels, within the waveform (the upper display) the excursions in sound pressure are large, meaning that air moves through the vocal tract with virtually no impedance, resulting in lots of sound energy and large excursions. In the spectrogram portion below that, we can see individual glottal pulses as vertical striations, and there is energy at all frequencies though there are some (lighter) bands of lesser energy and some (darker) bands of greater energy, especially corresponding to the formants of the specific vowel. The amplitude trace on top of the waveform confirms that the highest amplitudes are in the vowel portions. This is the canonical acoustic appearance of a very open vocal tract with voicing.

There is a change in amplitude and energy distribution as we move into the [d] portion of these words. The waveform excursions become smaller during [d], a change which happens relatively quickly. The amplitude display reduces during [d], and we also see that there is relatively little high-frequency energy in the spectrogram, where we see just a band of energy below 1000 Hz. This difference in appearance is because during production of [d], the vocal tract is constricted so that air does not flow out of the mouth, and there is a significant drop in acoustic energy. This vocal tract constriction also causes a buildup of air pressure behind it. The release of that constriction then results in a spike in acoustic energy at the transition from the stop closure to the following vowel, which we see as a sudden short darker period in the spectrogram – this is the "release burst". All of these abrupt and large drops and increases in energy are what tells us that [d] is a voiced stop.

Similar cues are found in the voiceless stop [t] in 'I sent' [ntʊmi] and 'now kill!' [ki:tɪ́].

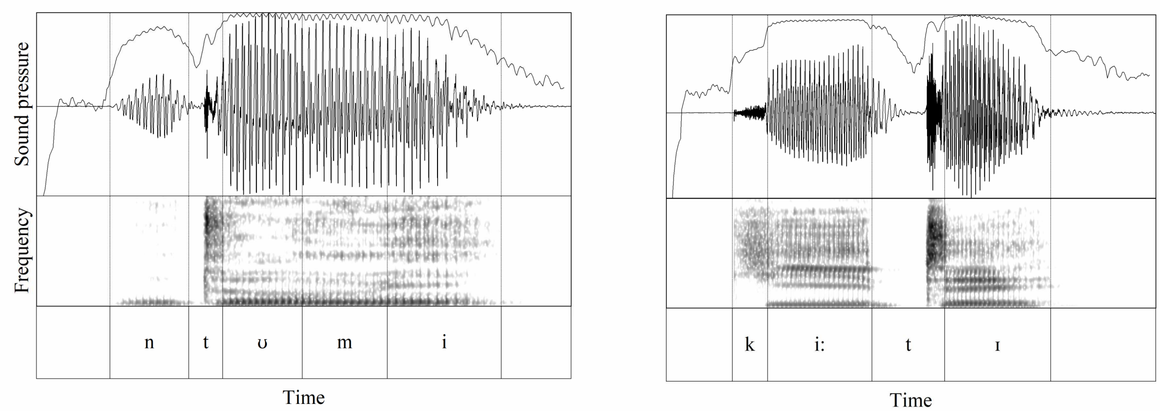

Figure 2

The main difference between [t] and [d] is that in [t] the vocal folds are wide open, so air flows at a very high rate, whereas in [d] the vocal folds are narrowed resulting in decreased airflow (owing to the fact that they are intermittently closed, about half of the time) although much acoustic energy results from the vocal folds banging together during [d] – that is what it means to be "voiced". In [ntʊmi], the closure in the oral cavity needed for [t] already exists during [n], so what happens is that the nasal passages close and the vocal folds become spread, causing a gradual decrease in acoustic energy until a point at the start of [t] where there is just a little energy. This nearly-silent period lasts up to about the midpoint of the portion marked as [t]. Then, suddenly, there is a short noise spike – the release burst of [t] – clearly visible in the waveform as well as in the spectrogram. After that there is a period of broad-spectrum noise until the point where the voiced vowel commences. In [ki:tɪ] on the right, we see a similar rapid change in the spectrogram from across-the-spectrum acoustic energy to just low frequency energy derived from residual vocal fold vibrations at the beginning of [t], vibrations which die out about halfway through the stop. After a period of virtual silence, the constriction is released and there is a substantial but brief release burst, then a transition period of about 25 msc as the vocal folds start to vibrate for the following vowel. Again, the rapid increase of acoustic energy when the stop is released is how we identify the sound as a stop.

The graphs in Figure 3 show [kʊzî:] 'we went' and [kʊ́vî:ta] 'to kill them'. These examples demonstrate the acoustic properties of non-stops, namely fricatives like [z] and approximants like [ʋ] (or [v], as spelled here).

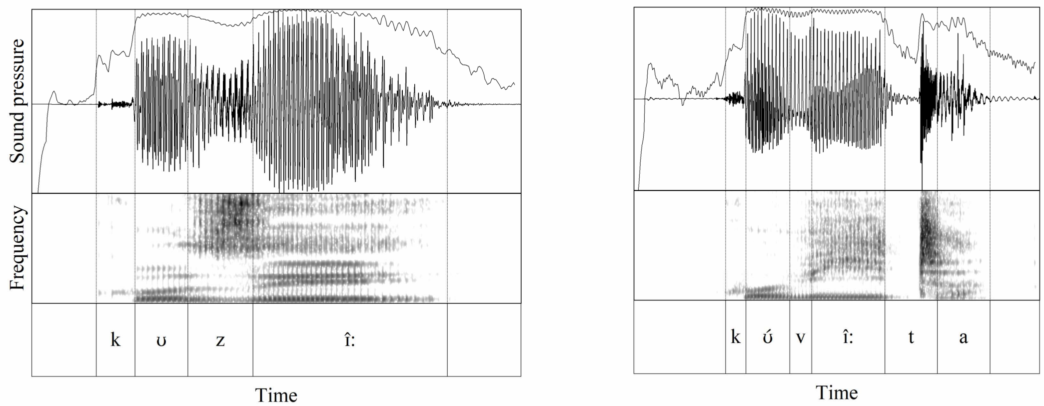

Figure 3

As we see from the spectrograms, waveforms and amplitude traces, [z] and [v] show a reduction in acoustic energy since there is some constriction during their production, but that reduction is much smaller compared to what is found in stops, it occurs more slowly, and there is no sudden change in energy corresponding to a release burst, because there is not a complete constriction and there is little increase in air pressure behind the narrowing of the airway where the consonant is articulated. It is difficult to assign precise edges to these consonants, and it would not be unreasonable to readjust the two boundaries of *z* and especially *v* a bit further to the left or right.

With this background on diagnosing the kind of constriction in consonants, we turn to what kind of consonant "g" is in Logoori (see g-variants online). The displays of [kʊgíní] 'we slandered' and [kʊgî:ka] 'to exaggerate' show canonical signs of [g] being a stop. There is a strong and rapid drop in acoustic energy in the spectrogram during *g*, a decrease in waveform excursions towards the end of *g*, and a sudden spike in acoustic energy at the very end of the consonant as *g* is released into the following vowel, all of which says that these are stops (and they sound like stops, listen to the online recordings). There is, however, a visible bit of difference within these two instances of *g*, which is that energy is more damped in the [g] part of the spectrogram of [kʊgíní] compared to [kʊgî:ka], and the release is longer in [kʊgî:ka] (above the voice bar at the bottom, there is more grey during [g] in [kʊgî:ka]).

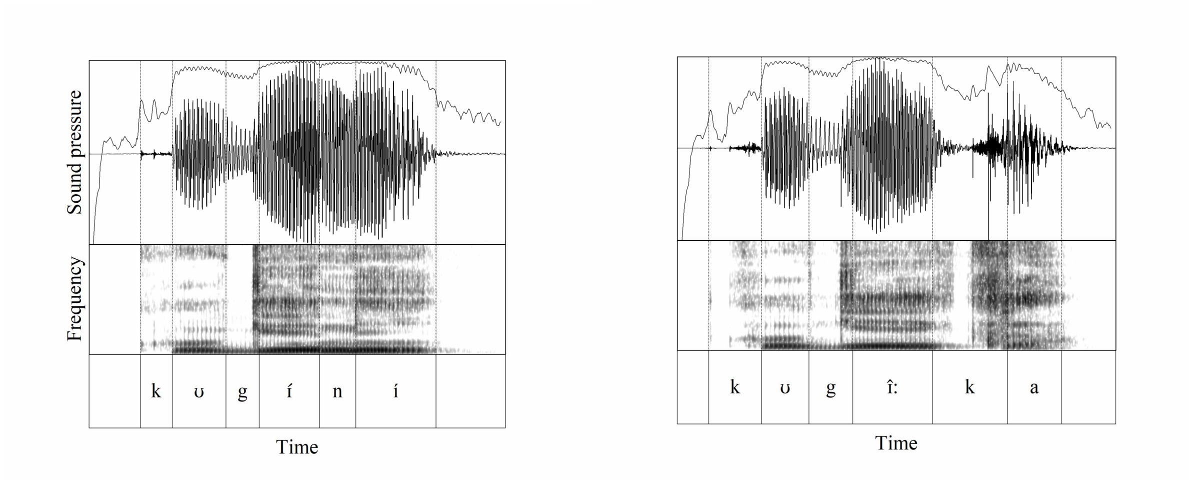

Figure 4

But the displays of [kɔgɛnja] 'to be amazed' and [kɔvɛ́ga] 'to shave' below present a very different picture of [g].

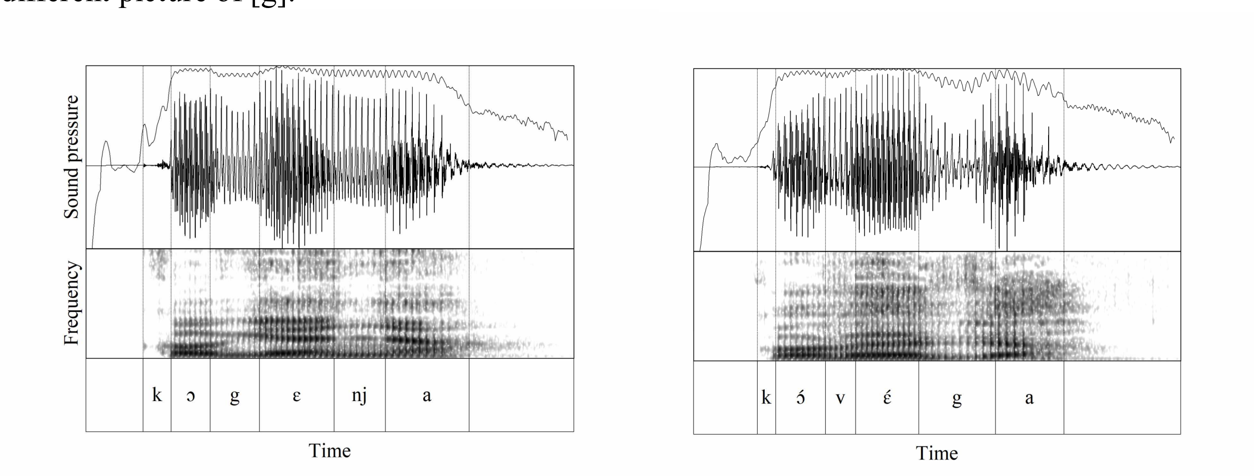

Figure 5

The waveform excursions here decrease going into and increase coming out of the consonant much less steeply, resulting in relatively little decrease in spectral energy compared to [kʊgíní] and [kʊgî:ka]. It is extremely hard to identify a specific time defining the beginning and end of *g* in these two examples – certainly there is no release burst. There are also differences in *g* between [kɔgɛnja] and [kɔvɛ́ga], where overall amplitude decreases more in [kɔvɛ́ga].

The displays of [kɔgɛnja] and [kɔvɛ́ga] are much more similar to [kʊzî:] and [kʊ́vî:ta], and as you can hear from the online recordings, they do not sound the same as the stop versions of *g*. It is possible that we have misidentified these consonants as being "the same thing", perhaps the language has both the stop [g] and the fricative [ɣ]. However, we would still face the problem that [kɔgɛnja] and [kɔvɛ́ga] differ from each other, as do [kʊgíní] and [kʊgî:ka]. Looking at yet another set of examples, [kʊgû:ta] 'to win', [kʊgʊti] 'we stacked', we again find more variation in how [g] is produced.

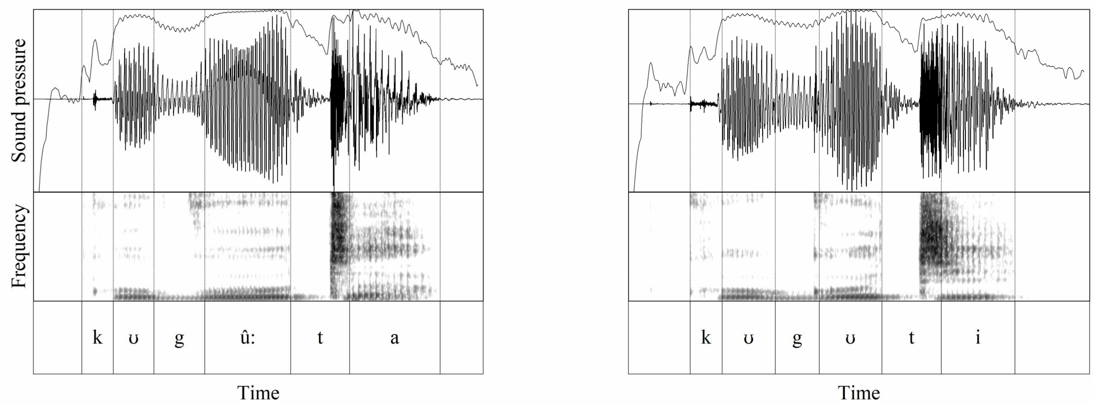

Figure 6

In [kʊgû:ta] and [kʊgʊti] looking at the spectrogram as well as the intensity contour, there is a general reduction in spectral energy, as one expects for stops, but these examples differ in the abruptness of the reduction in waveform excursions, which are more precipitous in [kʊgʊti]. The examples differ in that [kʊgʊti] has a brief release burst which is lacking in [kʊgû:ta]. What we have now seen is that in 6 examples of *g*, we have 6 different physical realizations of *g* in terms of the range of variation in its constriction – the prospects are good that if we add a few more examples, we will find even more variation.

The consonant [k] also exhibits considerable variability in degree of constriction during its production (follow online under *k*-variants to hear these words). In the first two examples [kɔkɛ:ŋga] 'to sculpt' and [vʊkî:ndʊ] 'cold wind' in Figure 7, we see reasonably clear examples of the full constriction that typifies stops. In both of these examples, there is strong damping of energy in the spectrogram at the beginning of the *k* interval. There is a very rapid decrease in waveform excursions at the start of *k* in [kɔkɛ:ŋga], but a more gradual decrease in those excursions in *k* of [vʊkî:ndʊ]. Then there is a sudden across-the-board increase in noisy spectral energy starting about at the middle of *k*, and there are weak signs of a release burst at that midpoint.

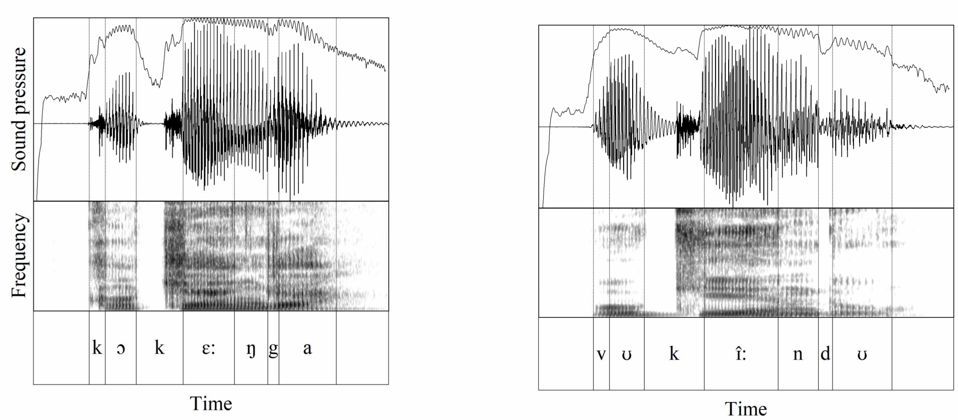

Figure 7

The main difference between these two words regards the dynamics of vocal fold abduction (not oral constriction), that in [kɔkɛ:ŋga] the vocal folds are quickly and strongly abducted but in [vʊkî:ndʊ], abduction is much slower.

In [[ɭíkâ:ŋga] 'guinea fowl' and [okoɭi] 'you did', we observe further variation. The decline of spectral energy at the beginning of the consonant is less rapid and less complete in these two recordings, compared to [kɔkɛ:ŋga] and [vʊkî:ndʊ], where there is not the same degree of suppression of higher frequency energy. We also see variation in the rate at which vocal fold vibration dies out at the left edge of the stop. Note also that [okoɭi] presents a double-tap release burst, a common feature of velar stops. These examples present a rather mixed picture.

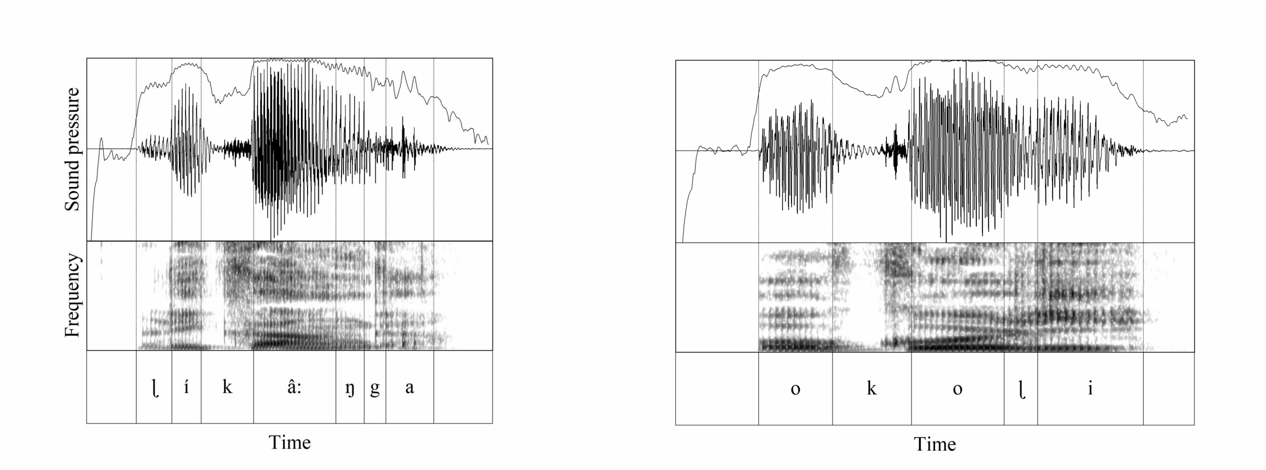

Figure 8

Our last two examples [kɪkʊ́nɛ] 'hamerkop' and [ndikî:gɪ] 'I would learn' show the other end of the constriction continuum. In neither case is there a release burst at the end of *k*. In [kɪkʊ́nɛ] the drop in spectral energy at the beginning of *k* is lesser and slower, and more takes the shape of a shift from voicing pulses to noise. In [ndikî:gɪ] there is not even the broad reduction in energy seen in [kɪkʊ́nɛ], instead we just have noise with a low-frequency "hole", very similar to what we saw above during [z] of [kʊzî:]. The constriction in [k] of [ndikî:gɪ] is so minimal that it better resembles the fricative [x], both in look and sound.

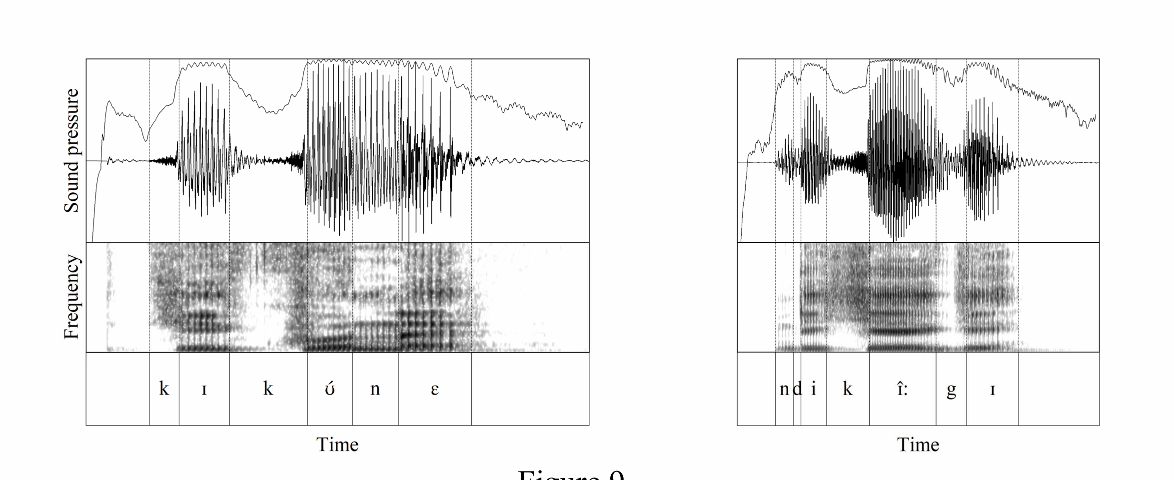

Figure 9

The overall picture which we see regarding [k, g] in Logoori is that the velar obstruents have a highly variable degree of constriction, which challenges any decision as to whether these sounds are "stops" or "fricatives". Phonological analysis into discrete categories presupposes that data fall into a small number of distinctive patterns, while phonetic analysis into a continuum means that there can be as many "patterns" as examples.

3.2. Types of variariation and phonological analysis

The situation with differences in Logoori velar constrictions is a canonical example of the notion of "phonetic implementation", which is the idea that every phonological unit generated in the grammar must be physically realised in a specific way – phonetically implemented. As has been repeated many times here and in other works, speech is highly variable. We have what seem to be the same thing (instances of *k*, or instances of *g*), which nevertheless are clearly different. The reality of speech is that "the same thing" from one perspective (a phonological analysis of speech into discrete sounds) can be "many different things", a continuum, perhaps in vocal fold abduction, perhaps in degree of oral cavity constriction. When one finds continuous variation in the physical measurements which defies segregation into two types, we suspect that we are dealing with phonetic variation, which requires more specialized acoustic, numeric analysis. Phonological variation is not about numeric values or acoustic properties, it is about category choices.

We have also seen (un-analyzed) examples of phonological variation in this chapter. Phonological variation is, as you would expect, the case where "a single thing" from a cerain perspective is realized as "two things" from another perspective. In the very first data set (1) (with corrections in (4) and (5)), we had the words [kʊdasa] 'to increase', [kɔgɛnja] 'to wonder', [kʊ́víta] 'to pass', [kɔl̥ɔ̂:ta] 'to dream'. Although these are different words, they are partially the same – they are verb infinitives, and they share a single marker for the infinitive. But the realized *form* of that marker varies between [kʊ] and [kɔ]. Since these words have this "same thing", our phonological puzzle is, why is the *form* of the prefix different in this set of words. Answering that question is the quintessential task of phonological analysis. We also see a tonal difference within the prefix between [kʊdasa] and [kʊ́víta], another phonological puzzle.

The answer in the vowel quality case is pretty simple. There is a phonological rule changing the vowel *ʊ* to *ɔ* when the vowel in the next syllable is either [ɔ] or [ɛ] – a change from one category of sound to another category, brought about by the particular category of sound following. We would then write a rule which accomplishes this, specifically we assume that the infinitive prefix is initially *kʊ*, but by applying that rule we can end up with the form [kɔ], which is then physically interpreted according to the phonetic rules of Logoori for producing [k], [ɔ] and so on.

Another example of variation in realization is that found within the single root for 'walk', the variation in vowels between [ɛ:] and [é:] as we have seen in [kɔgɛ:nda] 'to walk' versus [kogé:ndi] 'we walked'. In this case, the differences are of much greater physical and perceptual magnitude, and we can easily hear the difference which we took note of in our transcriptions, not requiring more specialized tools of acoustic analysis.

The primary decision which the linguist creating language data must make is which facts of the language should be represented as category choices, and which require phonetic analysis as a physical continuum. The kind of facts regarding voicing and degree of constriction of Logoori velars seen above fit squarely in the realm of "physical interpretation of the symbolic". Because data in grammatical descriptions of languages are presented in compact symbolic form (not detailed records of specific physical performances), grammars naturally present symbolic data – a phonological transcription – which stand for an open-ended class of utterances, covering both realized and potential sounds. Moreover, in the generative approach to language, these transcriptions do not stand for the physical behavior, they stand for the mental product underlying behavior, that is, the output of a phonological computation. It is not always obvious whether a particular fact of pronunciation is due to variability in production, or does it reflect computational options that change one sound category to another? We take up the matter of "contrast" as a diagnostic of phonological status in more detail in chapter 3.

Another kind of variation in language sound is "free variation", where an individual may freely choose between specific pronunciations of an utterance. For many speakers of English, the word "electric" can be pronounced either as [əˈlɛktrɪk] or [iˈlɛktrɪk]. FI can pronounce "he drove" either as [ahiɾi] or [aʃiɾi], and 'I cut' as [ŋkaɾi] or [ŋgaɾi] – but he only says [akaɾi] 'he cut', never *[agaɾi]. This is a matter for phonological analysis, because there are two possible pronunciations coming from one and the same mental grammar, and phonological analysis is *at least* about what an individual grammar produces. In fact, he is aware of the two variants, can freely produce either, and has opinions about which one he prefers. A speaker's ability to consciously identify one versus another class of pronunciations is very strong evidence that we are dealing in cognitively distinctive categories.

Alongside individual free variation there is also "dialect variation", where some individuals pronounce words and phrases one way but other individuals pronounce them a different way. I pronounce the words "cot" and "caught" the same, but some speakers pronounce "cot" as [kat] but "caught" as [kɔt]. I pronounce the negative form of [kæn] "can" as [kæ̃ʔ], but many speakers pronounce it as [kæ̃nt]. FI pronounces the Logoori word for 'goat' as [ɪmbúɭi] whereas another speaker, Editon Mulera, pronounces it as [ɪmbúɾi] – though both say [ɪmíɾi] for 'swallow'. The two individuals differ in age by a generation and come from opposite areas of Logoori-speaking territory, so it is unsurprising for there to be dialect variation. The issue which the grammar-author must address is whether to report this variation at all (rather than just giving a presumed standard pronunciation), and if so, *how* is this information to be reported? It is usually unclear to what extent a speaker of a language is aware of dialectal variation in pronunciation in the language and therefore whether dialectal relations play a role in mental grammars. Having learned of the pronunciation difference between “cot” and “caught” in introductory linguistics 50 years ago, I still do not know which words use [a] and which use [ɔ] although I have developed some guessing strategies. In fact, even for speakers who have both vowels there are substantial differences between speakers in terms of what word has which vowel. On the other hand, FI’s knowledge of the rule – used by other speakers – where his *nt,ŋk* always becomes *nd, ŋg* is quite reliable and easy to compute, so it *could* be treated as an optional rule in his dialect.

The most general descriptive practice is to report the facts of “a language”, considered in a broad social sense. For well-studied major national languages, the author will typically focus on English, Arabic, Italian, Chinese, Turkish or Finnish, meaning the standard written varieties. In such cases, data may not even be derived from listening to a speaker and transcribing what was uttered, they data may instead come from a standard pronouncing dictionary, which smoothes out observed speech variation and favors a specific variant (e.g. “RP” in British English, Parisian French in France, Hanover dialect in Germany, Beijing Mandarin in China). Even in grammars based exclusively on transcribed speech, facts of language performance typically aim to report dominant trends rather than documenting all variations, which means that many facts of pronunciation, even important phonological differences, will not be reported. Because there is a wide range of variation in linguistic descriptive practices, more or less information may be included in a grammar, depending on the author’s interpretation of what counts as phonologically significant, or what version of the language is being presented.

The ideal descriptive grammar, for phonological purposes, is a paradigm-oriented statement of the structure of the language. This chapter has presented some of the building blocks for a paradigmatic grammar of Logoori, i.e. verb infinitives, past tense with 1pl subject, a few samples of imperatives in the singular and plural, and further data from the language used in this book will enable a somewhat deeper understanding of the language. On the other hand, some grammars are more focused on explaining the communicative function of phrasal combinations of words, and may gloss over details of how the shape of a morpheme varies as a function of what elements the morpheme combines with. Logoori, which has a very rich and productive morphology and phonology, has the potential for creating about a hundred thousand variants of a single verb root, so listing “all of the verb forms” in the language is impossible. The concept of “grammatical rule” is thus central to understanding the structure of languages like Logoori. The purpose of this book and course is to lay out the analytic methodology of describing the phonological grammar of languages.

3.3. Phonological categories

Our transition from physical speech to mental units begins by converting perceived speech into sequences of styandard written symbols representing the different mental symbols in the phonology of a language. The conversion is primarily guided by the existence of an auditory standard, in the form of the International Phonetic Alphabet, which pairs a particular physical description (“close-mid front rounded vowel”) with auditory exemplars. That chart (above and online) also embodies a claim as to granularity, in that the accepted standard includes a difference between [i, ɪ, e] but has no specific symbol that represents how “i” is pronounced in Turkish, nor does it represent the myriad physical vowels differences in nominally-transcribed [e, ɛ] in Kikuyu, Kamba, Meru, Logoori, Sukuma, Matumbi, Sotho – just to mention a handful of Bantu languages. There are adjustment-diacritics allowing one to indicate in what direction a particular sound deviates from the standard of the symbol, but the set of auditory samples for applying these adjustments are very limited. The implicit theory underlying the IPA assumes that only certain aspects of speech are linguistically relevant, but does not entail any particular granularity, so the Handbook of the IPA does not address cross-linguistic differences in [i], [e], or [ʕ]. However it is noted that the symbol [β], defined as a “bilabial fricative”, can be used to represent a “bilabial approximant” – in fact, [ʋ] which is defined as a “labiodental approximant” is often used to represent a bilabial approximant. This reflects an implicit limit on the what is linguistically important in human languages – we know of no language compelling us to distinguish bilabial and labiodental approximants. When faced with sufficient evidence that a sound not included in the IPA chart is an “important distinction”, as in the case of [ⱱ], the Council of the IPA is willing to modify the chart.

The IPA chart divides sounds into a few cross-cutting categories such as “bilabial”, “labiodental”, “postalveolar”, “fricative”, “plosive” where the terms refer to aspects of their phonetic realization. Phonological usage adheres to a somewhat different phonologically-motivated classification and terminology, for example “plosives” are termed “stops” in phonological usage, “close” vowel are “high” in phonological usage, and “postalveolar” is “alveopalatal” in phonological usage. Phonology makes use of a phonologically-relevant concept “liquid”, which has no analog in the IPA chart, the same goes for the phonological term “glide”, which is a subset of the IPA “approximants”. More to the point, as we discuss in Chapter 4, terms like “bilabial” and “tap” have no formal status in phonology, they are just convenient ways of talking about a particular sound. That chapter systematically presents standard phonological terms and their ontological status – as “features”.

3.4. Transcriptional variation

Given the preceding explanation of the nature of transcriptions, one would expect that there should only be a single transcription of a given English word. Observationally, this is false, there is massive variation in transcriptions of English words. One obvious source of variation is dialect difference, the fact that a speaker raised in the Pacific Northwest of the US actually talk different from someone raised in Newcastle UK, Auckland NZ, Capetown SA, Nairobi KE or even Vancouver BC. Even if we stick to a single dialect, you can encounter a lot of variation in transcriptions, for example the word *great* may have [k g g̊] as the first consonant, [r ɾ ɹ ɹʷ] as the second consonant, [e eː ei ɛi eɪ ɛɪ ej ɛj] as the vowel and [t t̚ tˀ ʔ] as the final consonant. The differences do not come from massive differences in pronunciation, they come from different ideas as to what is important to represent in a phonological transcription. The variation [k g g̊] reflects the fact that word-initial *g* in English is not pronounced the same as it is in French, with clear vocal fold vibration. The phonologically-mandatory fact that has to be recorded is that "great" and "crate" are different words, pronounced differently. Marking this difference as [k] versus [g] is one way to do it, so is [k^{h}] versus [k], or [k] versus [g̊] (meaning "g, without vocal fold vibration. The letter *r* would be pronounced [ɹ^{w}] in my speech, but many people from the same dialect area pronounce the consonant as [ɹ], and it is not linguistically important to be overly specific about exact pronunciation. It is also much more efficient to type [k, g] rather than [k^{h}, g̊] when transcribing English words. The variation [t t̚ t^{ʔ} ʔ] probably attends too closely to the pronunciation of specific utterances, these are all a single consonant at the end of a word, pronounced different ways. The vowel differences generally reflect different ideologies regarding the phonological status of "long" English vowels, although there *is* also a real pronunciation difference between dialects where "great" has a tense monophthong and those where it is a diphthong of some sort. The phonologically most-relevant question for transcription is, how do you symbolize the *difference* between "gate" and "get".

It is with this understanding of what phonological data are and where they come from that we now proceed to methods of phonological analysis in Chapter 3.