Introductory Phonology

Department of Linguistics, UCLA

Contents

Introductory Phonology

Preface

This text is meant as a first course book in phonology. The book has evolved as the textbook for a course taught to a mostly undergraduate audience over a number of years in the Department of Linguistics at UCLA. The course meets in lecture for four hours per week, with a one hour problem-solving session, during a ten-week term.

The ideal audience for this book is a student who has studied some linguistics before (and thus has some idea of what linguists are trying to accomplish), and has already taken a course in general phonetics, covering at least the basics of articulatory phonetics and the International Phonetic Alphabet. It is possible to make up this material on the fly through reading and practice,[1] but I consider this strategy second-best. A short chapter on phonetics, intended for review, is included in this text.

As the title implies, this book is meant to be an introductory text. By this I mean not that it is meant to be easier than other texts, but rather that it emphasizes the following two things:

- **Analysis of phonological data**, along with methods that experience has shown can be useful in leading to accurate analyses.
- The **scientific context** of phonological analysis: what are we trying to understand when we carry out formal analyses of the phonological patterns of languages?

I consider the first item to be crucial in an introductory course, because if analysis is not well done at a basic level, all of the more sophisticated theoretical conclusions that might be drawn from it become untrustable. The second item is likewise crucial, to make phonological analysis meaningful.

As a consequence of these general goals, I have left out quite a few topics that currently are of great interest to many phonologists, myself included. This reflects my goal of teaching first the material that will provide the most solid foundation for more advanced theoretical study.[2]

I have tried to avoid a common problem of linguistics textbooks, that of presenting data simplified for pedagogical purposes without providing some means for the student to access more

[1] Some recommended material for this purpose: *A Course in Phonetics* by Peter Ladefoged (5th ed., 2005, Heinle), and the accompanying sound materials made available at http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants.

[2] For students going on to more advanced topics, I have found the following texts to be helpful: John Goldsmith (1990) *Autosegmental and Metrical Phonology* (Oxford: Basil Blackwell); Michael Kenstowicz (1994) *Phonology in Generative Grammar* (Oxford: Basil Blackwell); René Kager (1999) *Optimality Theory* (Cambridge: Cambridge University Press); John McCarthy (2002) *A Thematic Guide to Optimality Theory* (Blackwell, Oxford).

information about the language. This is provided in the "Further Reading" sections at the end of each chapter.

A number of passages in the text offer guidance in eliciting useful and valid data from native speakers. This relates to the phonology course I teach, in which one of the major assignments is a term paper involving analysis of data gathered first hand from a native speaker.

A computer resource for phonology that I have found useful in conjunction with this text is **UCLA FeaturePad**, a computer program created by Kie Zuraw, which helps students to learn and use features by showing the natural classes that correspond to any selection of feature values. It also shows how the sounds are changed when any feature values are changed. The program may be downloaded for free from http://www.linguistics.ucla.edu/people/hayes/120a/.

Many people provided me with help and feedback on this text, for which I am very grateful. Among them were Marco Baroni, Christine Bartels, Roger Billerey, Abigail Cohn, Maria Gouskova, Patricia Keating, Charles Kisseberth, Jongho Jun, Sun-Ah Jun, the late Peter Ladefoged, Lisa Lavoie, Margaret MacEachern, Donka Minkova, Susan Moskwa, Pamela Munro, Russell Schuh, Shabnam Shademan, Bernard Tranel, Adam Ussishkin, Keli Vaughan, and Kie Zuraw. I'm certain that I've left names out here, and in cases where my memory has failed me I hope the unthanked person will understand.

I welcome comments and error corrections concerning this text, which may be sent to bhayes@humnet.ucla.edu or Department of Linguistics, UCLA, Los Angeles, CA 90095-1543.

Portions of chapters 2, 3, 6, and 7 appeared in earlier form as Chapter 12 of *Linguistics: An Introduction to Linguistic Theory* by Victoria Fromkin et al., (2000, Blackwell).

Los Angeles
July 2007

Chapter 1: Phonetics

1. Phonetics and phonology

There are two branches of linguistic science that deal with speech sounds: **phonetics** and **phonology**.

Phonetics is primarily an experimental science, which studies speech sounds from three viewpoints:

- **Production**: how sounds are made in the human vocal tract
- **Acoustics**: the study of the waveforms by which speech is transmitted through the atmosphere
- **Perception**: how the incoming acoustic signal is processed to detect the sound sequence originally intended by the speaker

Phonology is also, sometimes, an experimental science, though it also involves a fair degree of formal analysis and abstract theorizing. The primary data on which phonological theory rests are phonetic data, that is, observations of the phonetic form of utterances. The goal of phonology is to understand the tacit system of rules that the speaker uses in apprehending and manipulating the sounds of her language (more on this in chapter 2).

Since phonological data are phonetic, and since (as we will see) the very nature of phonological rules depends on phonetics, it is appropriate for beginning students to study phonetics first. In particular, a phonologist who tries to elicit data from native speakers without prior training in the production and perception of speech sounds will be likely to have a hard time. The material that follows can be taken to be a quick review of phonetics, or else a very quick introduction that can be amplified with reading and practical training from materials such as those listed at the end of the chapter.

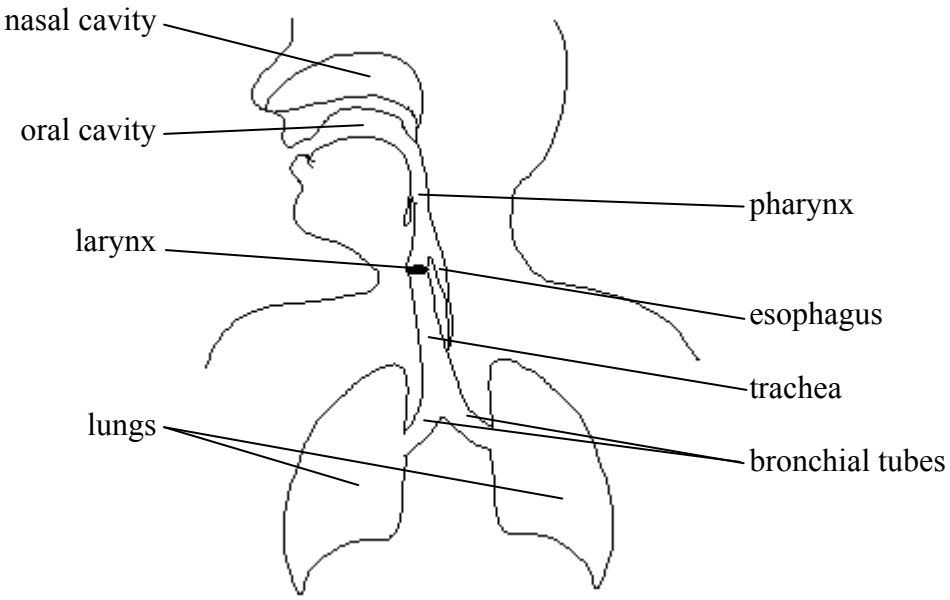
In principle, a phonologist should understand all three of the areas of phonetics listed above: production, acoustics, and perception. Of these, production probably has the greatest practical importance for the study of phonology. Since it is also the simplest to describe, it is what will be covered here.

2. The vocal tract

The term "vocal tract" designates all the portions of the human anatomy through which air flows in the course of speech production. These include (from bottom to top):

- The lungs and lower respiratory passages
- The **larynx** (colloquially: "voice box"). This is the primary (but not the only) source of sound in speech production

- The passages above the larynx, called the **pharynx**, **oral cavity** (mouth), and **nasal cavity**

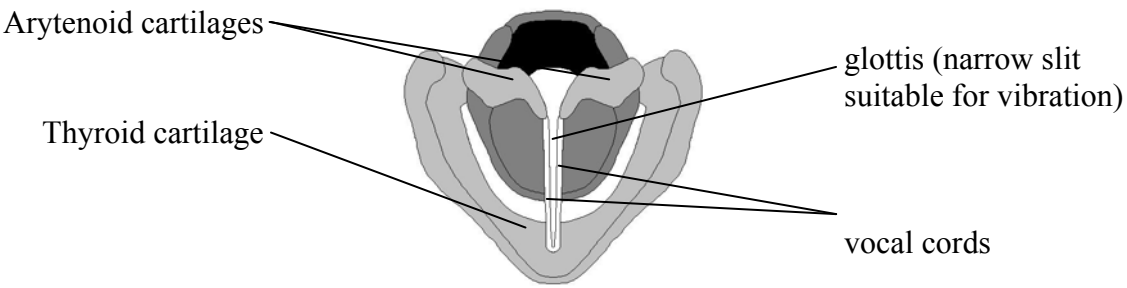


The lungs and respiratory muscles produce a fairly steady level of air pressure, which powers the creation of sound. There are occasional momentary peaks of pressure for certain speech sounds and for emphatically stressed syllables. Air from the lungs ascends through the **bronchial tubes**, which join to form the **trachea** (windpipe). The bronchial tubes and the trachea form an inverted Y-shape.

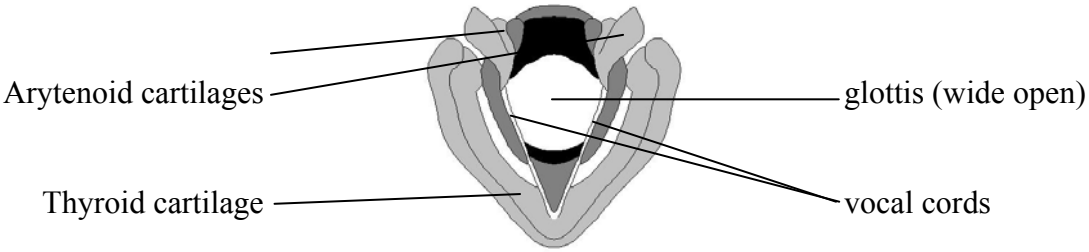
2.1 The larynx

The larynx is a complex structure of cartilage and muscle, located in the neck and partly visible in adult males (whose larynxes are the largest) as the "Adam's apple." Here are two diagrams of the larynx:

Larynx with vocal cords in position to vibrate:



Open larynx (vocal cords spread):



The larynx contains the **vocal cords** (not "chords"), which are parallel flaps of tissue extending from each side of the interior larynx wall. The vocal cords have a slit between them, called the **glottis**. The vocal cords are held at their rear ends by two small cartilages called the **arytenoid cartilages**. Since these cartilages are mobile, they can be used to adjust the distance between the vocal cords.

When the vocal cords are held tightly together, the sound known as a **glottal stop** is produced; it can be heard in the middle of the expression "uh-oh" and is used as a speech sound in many languages.

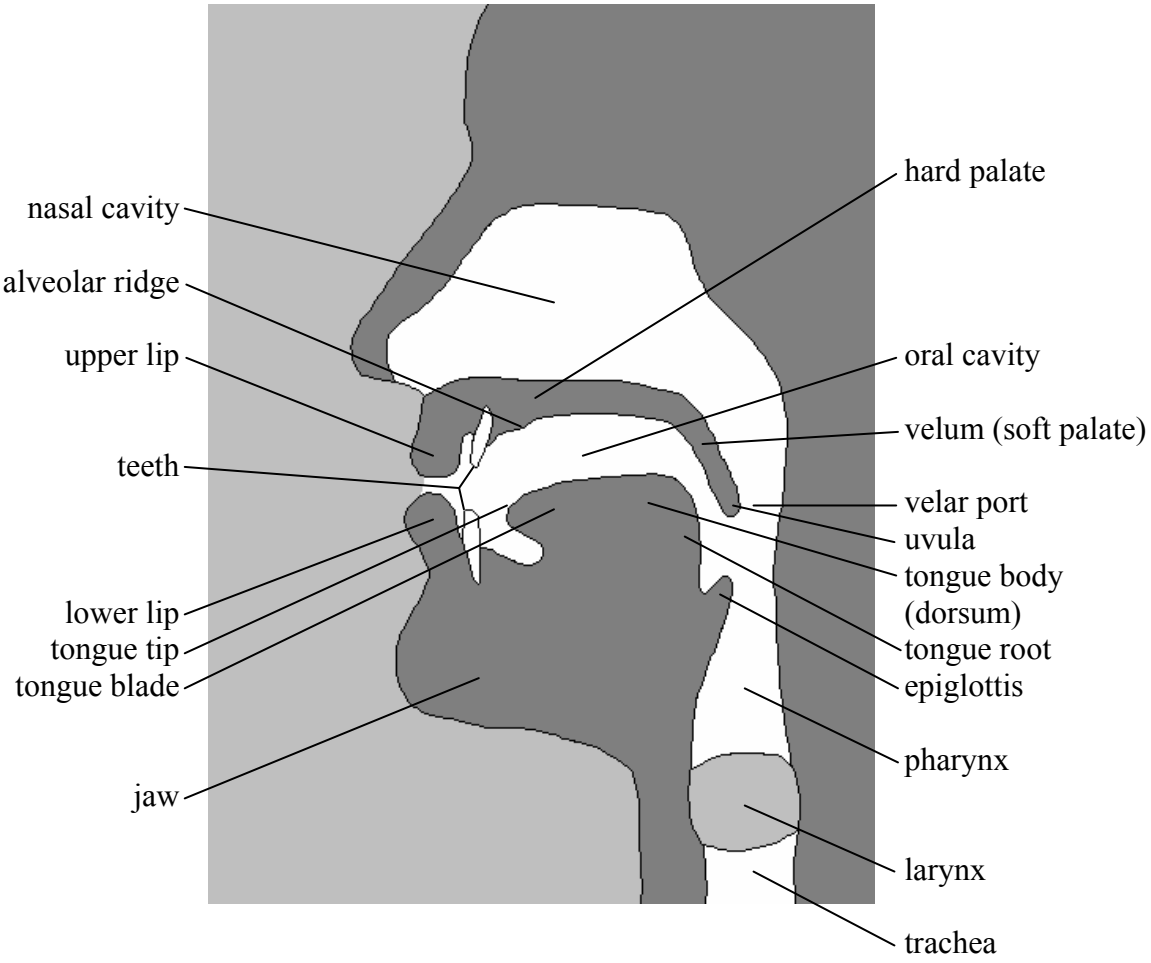
If the vocal cords are placed close to each other but not tightly shut, and there is sufficient airflow from the lungs, then the vocal cords will vibrate, creating **voicing**. This is the configuration shown in the first diagram above. Voicing is the most important and noticeable sound source in speech.

The vocal cords can also be spread somewhat apart, so that air passing through the glottis creates turbulent noise. This is the way an "h" sound is produced. The vocal cords are spread farther still for normal breathing, in which airflow through the larynx is smooth and silent. This is the configuration shown in the second figure above.

The cartilages of the larynx, especially the **thyroid cartilage** to which the front ends of the vocal cords attach, can stretch and slacken the vocal cords, thus raising and lowering the pitch of the voice. This is somewhat analogous to the changes in pitch that occur when a guitar string is tightened or loosened.

2.2 The upper vocal tract

Sound created at the larynx is modified and filtered as it passes through the upper vocal tract. This area is the most complex and needs the most detailed discussion; you should refer to the diagram below while reading the text.



The main route through the upper vocal tract is a kind of arch, starting vertically upward from the larynx and bending forward through the mouth. There is an opening about half way from larynx to lips, called the **velar port**, through which air can pass into the nasal passage and outward through the nostrils. In the diagram above, the velar port is wide open.

We will first cover the upper surface of the upper vocal tract (the roof of the mouth and the back of the pharynx), then the lower surface (floor of mouth, continued as the front wall of the pharynx).

Going in the "upstream" direction, the crucial landmarks of the upper surface are:

- The upper **lip**.
- The upper **teeth** (in particular, the incisors).

- The **alveolar ridge**, a bony ridge just behind the base of the upper incisors. Most people can feel their alveolar ridge by moving the tongue along the roof of the mouth.[3]
- The **hard palate**, which is that part of the roof of the mouth underlain by bone. You can feel the hard palate, and its rear edge, with the tip of your tongue
- The **velum**, or **soft palate**. This is a flap of soft tissue that separates the mouth from the nasal passages. It is attached at the front (to the hard palate) and at the sides, but hangs loose at its rear edge. Various muscles can raise and lower the velum. When the velum is high, then the velar port is closed, and air is confined to the oral passage.[4]
- The little thing that dangles from the rear edge of the velum is called the **uvula** ([ˈjuːvjələ]), Latin for "little grape". The uvula is vibrated (trilled) as a speech sound in some languages.
- Once we are past the velum, we are no longer in the mouth proper but in the rearward part of the upper vocal tract, commonly called the **pharynx**. The rear pharyngeal wall is continuous and has no significant landmarks all the way down to the larynx.

The crucial parts of the lower surface of the upper vocal tract are as follows:

- The lower lip and the tongue rest on the **jaw**, which raises and lowers the lower lip and tongue when it moves during speech.
- The **lower lip** is more mobile than the upper in speaking, though both move considerably. They can touch one another, closing the mouth, or the corners of the lips can be pulled in, creating lip rounding.
- The **tongue** is somewhat deceptive in its size and shape. The parts that are obvious to an external observer are the **tip** (sometimes called the **apex**) and the **blade**. These are merely an appendage to the much larger tongue **body** (also called **dorsum**), a roundish muscular body that can move in all directions. Movements of the dorsum can radically change the shape of the vocal tract, a fact that is crucial in the production of distinct vowel sounds.
- The rear surface of the dorsum is called the **tongue root**. Behind it is a flap called the **epiglottis**.

3. Describing speech sounds

The human vocal tract can produce thousands of audibly distinct sounds. Of these, only a subset are actually used in human languages. Moreover, of this subset, some sounds are much more common than others. For example, almost every language has a *t*-like sound, whereas very

[3] Some people do not have a sharply defined alveolar ridge.

[4] If you can produce a distinction between nasal and oral vowels, as in French or Portuguese, then it is possible to watch the velum work, using a flashlight and a mirror. When a speaker alternates between oral and nasal vowels, the velum is seen to billow up and down like a sail (which is what "velum" literally means in Latin).

few languages have an epiglottal stop or a bilabial trill. Any one language uses only a fairly small inventory of distinct speech sounds, usually no more than a few dozen.

The commonsense distinction of **vowels** and **consonants** is a generally valid one for phonetics and phonology. Roughly, vowels are highly sonorous sounds, made with a relatively open vocal tract. Consonants involve some kind of constriction (or more than one constriction) in the vocal tract. They are quieter than vowels, and often are detectable by the ear not so much by their own sound as by the transitional acoustic events that occur at the boundary of consonants and vowels.

For both vowels and consonants, phonetic description involves assigning a **phonetic symbol** to each sound. This book will use the standard, internationally-accepted phonetic symbol set called the International Phonetic Alphabet (IPA), promulgated by the International Phonetic Association.

It should be clear why the use of standard symbols, rather than spelling, is crucial. The spelling systems of most languages are ambiguous (consider *read*, *bow*) and inconsistent in the depiction of identical sounds (consider *whole*/*hole*, *real*/*reel*). Cross-linguistically, the situation is even worse, as different languages employ the same letters to depict different sounds: the letter *j* spells four quite different sounds in English, French, Spanish, and German (in IPA these are [d͡ʒ], [ʒ], [x], and [j]). Since this book will be presenting data from many languages, I will standardize all data using IPA transcription.

Phonetic transcription is traditionally given surrounded by square brackets. Thus, one possible rendition of the previous sentence in IPA (as pronounced in my own dialect of English) is:

[fəˈnɛɾɨk t͡ʂɹæ̃nˈskɹɪpʃə̃n ɨz t͡ʂɹəˈdɪʃɨnəli ˈgɪvə̃n səˈɹæ̃ũndɨd ba͡ɪ ˈbɹækɨts]

4. Consonants

Consonants are classified along three dimensions: **voicing**, **place of articulation**, and **manner of articulation**.

4.1 Voicing

In a voiced consonant, the vocal cords vibrate. For example, the "s" sound, for which the IPA symbol is simply [s], is voiceless, whereas the "z" sound (IPA [z]) is voiced. If you say "sa, za" while planting the palm of your hand firmly on the top of your head, you should feel the vibrations for [z] but not for [s].

The sounds [p t k] are voiceless. The sounds [b d g] as they occur in (for example) French or Japanese are voiced; in English they are often voiced for only part of their duration or even not at all; nevertheless the symbols [b d g] are traditionally used for them.

4.2 Manner

There are various manners of articulation.

In a **stop**, the airflow through the mouth is momentarily closed off. This can be done by the two lips, forming [p] or [b]; by the tongue tip touching the alveolar ridge, forming [t] or [d]; by the tongue body touching the palate, forming [k] or [g], and in other ways.

In a **fricative**, a tight constriction is made, so that air passing through the constriction flows turbulently, making a hissing noise. Some of the fricatives of English are [f], [v], [θ] (the first sound of *thin*), and [ð] (the first sound of *the*). In **sibilant** fricatives, the mechanism of production is more complex: a stream of air is directed at the upper teeth, creating noisy turbulent flow. The four sibilant fricatives of English are [s], [z], [ʃ] (the first sound of *shin*), and [ʒ] (the consonant spelled *s* in *pleasure*).

An **affricate** is a stop followed by a fricative, made at the same location in the mouth in rapid succession so that the result has the typical duration of a single speech sound. English has two affricates: voiceless [t͡ʃ] (as in *church*) and voiced [d͡ʒ] (as in *judge*). As can be seen, the IPA symbol for an affricate is made with the symbols for the appropriate stop or fricative, optionally joined with a ligature.

Affricates are often considered to be a species of stop; that is, "affricated stops."

In a **nasal** consonant, the velum is lowered, allowing air to escape through the nose. Most nasal consonants have a complete blockage within the mouth at the same time. The places of articulation for nasals are mostly the same as those for stops. The nasal consonants of English are [m] (*Mom*), [n] (*none*), and [ŋ] (*young*).

Nasals like [m, n, ŋ] in a certain sense are also stops, since they involve complete closure in the mouth; hence the term "stop" is ambiguous. I will use this term here in its strict sense, which includes oral stops only.

In **taps** and **flaps**, an articulator makes a rapid brush against some articulatory surface. The motion of articulator is forward in a flap, backward in a tap. North American varieties of English have alveolar taps (IPA [ɾ]) in words like *lighter* and *rider*.

In a **trill**, an articulator is made to vibrate by placing it near an articulatory surface and letting air flow through the gap. Many dialects of Spanish have an alveolar trill (IPA [r]) in words like *perro* 'dog'. The uvula ([ʀ]) and lips ([ʙ]) can also be trilled.

Approximants are consonants in which the constriction is fairly wide, so that air passes through without creating turbulence or trilling. In **lateral approximants**, the air passes around the sides of the tongue, as in English [l]. In **central approximants**, the flow is through a gap in

the center. English dialects have (at least) three central approximants, namely [j],[5] as in *youth*, [w], as in *win*, and [ɹ], as in *roar*.

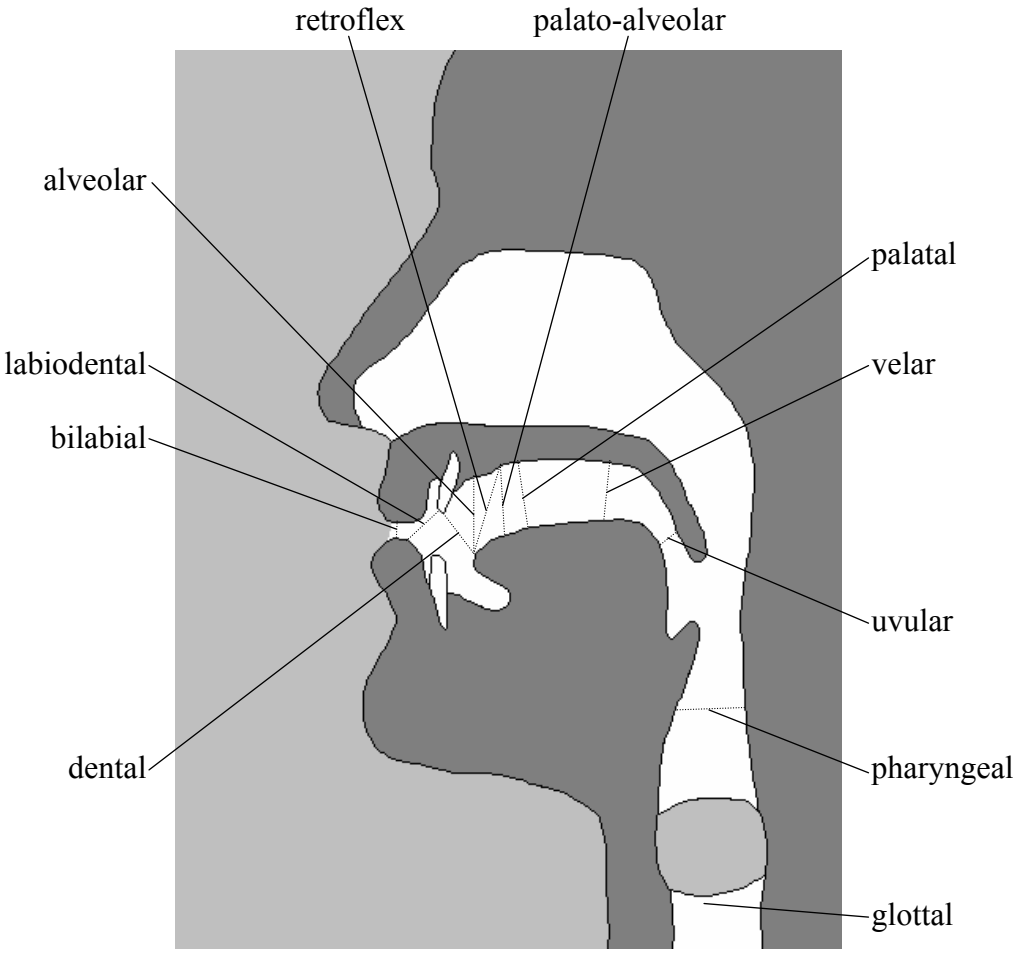
The last three categories just given are sometimes presented with a different classification. The **liquids** are the sounds that have the characteristic acoustic quality of l-like and r-like sounds.[6] This term groups [l] and similar sounds together with tap [ɾ], trilled [r], approximant [ɹ], and various similar r-sounds. Under this same scheme, the **glides (**also called **semivowels**) are the central approximants; that is, [j], [w], and similar sounds covered below.

4.3 Place of articulation

I will cover most of the possible places of articulation, proceeding from front to back. Each place is shown on the following diagram. In the diagram, dotted lines indicate the approximate path taken by an articulator in making contact with the opposite wall of the vocal tract.

[5] The IPA symbol is modeled after the spelling of German, Dutch, Polish, Swedish, and various other languages. The letter *y*, used in English, has a different meaning in IPA, given below.

[6] In terms of speech acoustics, *l*-like sounds have an exceptionally high third **formant** (band of acoustic energy), and *r*-like sounds have an exceptionally low third formant. Non-liquid approximants have third formants that would be expected, given their first and second formants.



- **Bilabial** sounds are made by touching the upper and lower lips together. English has a voiceless bilabial stop [p], a voiced bilabial stop [b], and a (voiced) bilabial nasal [m].

  Note that the description just given follows the standard form for describing a consonant: **voicing**, then **place**, then **manner**. In the case of nasals and approximants, which are normally voiced, it is common to specify only place and manner.

- **Labio-dental** sounds are made by touching the lower lip to the upper teeth. English has a voiceless labio-dental fricative [f] and a voiced one, [v].

- **Dental** sounds are made by touching the tongue to the upper teeth. This can be done in a number of ways. If the tongue is stuck out beyond the teeth, the sound is called an **interdental**, though we will not be concerned with so fine a distinction. English has a voiceless dental fricative [θ] (*th*in), and a voiced one [ð] (*th*e).

- **Alveolar** sounds are made by touching the tip or blade of the tongue to a location just forward of the alveolar ridge. English has a voiceless alveolar stop [t], a voiced alveolar stop [d], voiceless and voiced alveolar fricatives [s] and [z] (both of them sibilants), a

voiced alveolar nasal [n], a voiced alveolar lateral approximant [l],[7] and a voiced alveolar central approximant [ɹ].

- **Palato-alveolar** sounds (sometimes called **post-alveolar**) are made by touching the blade of the tongue to a location just behind the alveolar ridge. English has a voiceless palato-alveolar fricative [ʃ] (*sh*oe), a voiced palato-alveolar fricative [ʒ] (*vi*s*i*on), a voiceless palato-alveolar affricate [t͡ʃ], (*ch*ur*ch*), and a voiced palato-alveolar affricate [d͡ʒ] (*j*u*dg*e).

- **Retroflex** sounds are made by curling the tongue tip backward, and touching the area just behind the alveolar ridge. Some English speakers lack the alveolar approximant [ɹ] and instead have a retroflex one, transcribed [ɻ]; retroflex stops and affricates are common in languages of India and Australia.

  In the strict sense of the term, palato-alveolars and retroflexes have the same place of articulation: the same place on the roof of the mouth is contacted. They differ in the part of the tongue (blade or tip) that makes the contact. Conventionally, however, palato-alveolar and retroflex are referred to as separate places of articulation.

- **Palatal** sounds are made by touching the tongue blade and the forward part of the tongue body to the hard palate. [j] (*y*oung) is sometimes described as a palatal approximant (see §5.5 below for a different kind of description); various languages have a variety of other manners of articulation at the palatal place.

- **Velar** sounds are made by touching the body of the tongue to the hard or soft palate. English has three velar sounds: a voiceless velar stop [k], a voiced velar stop [g], and a velar nasal [ŋ] (si*ng*).

- **Uvular** sounds are made by moving the tongue body straight back to touch the uvula and neighboring portions of the soft palate. The "r" sound of French and German is usually a voiced uvular fricative, [ʁ]. The nasal consonant that occurs at the end of many words in Japanese is often pronounced as uvular [ɴ].

- **Pharyngeal** sounds are made by moving the tongue body down and back into the pharynx. A voiceless pharyngeal fricative is transcribed [ħ]; it occurs for example in Arabic.

- **Glottal** sounds are made by moving the vocal cords close to one another. English has a voiceless glottal fricative [h].

4.4 Consonant chart

Below is reproduced the part of the official IPA chart covering consonants.

[7] For many speakers of English, the *l* sound is actually dental [l̪].

CONSONANTS (PULMONIC)

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar [=palato-alveolar]	Retro-flex	Palatal	Velar	Uvular	Pharyn-geal	Glottal
Plosive [= stop]	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
[Central] approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

It can be seen that any consonant in the chart is describable with the terminology given above, and that a fair number of sounds are listed that do not occur in English. Quite a few of these will come up in the chapters to follow.

The symbols for dentals, alveolars, and palato-alveolars are systematically ambiguous. Where it is important to make a distinction, it is possible to do so with diacritics:

- [t̪] = voiceless dental stop
- [t] = voiceless alveolar stop
- [t̠] = voiceless palato-alveolar (= postalveolar) stop

Affricates are formed by joining a stop and fricative symbol together, as in for instance [t͡ʃ], the first and last sound of *ch*urc*h*. The same ligature may be used for so-called "complex segments" such as labial-velar [k͡p], which are formed at two places of articulation simultaneously.

A subsidiary part of the IPA chart covers consonants in which the airflow comes not from the lungs, but from motions of the larynx (**implosives**, with inward airflow; and **ejectives**, with outward); or of the tongue body (**clicks**).

CONSONANTS (NON-PULMONIC)

Clicks		Voiced implosives		Ejectives	
ʘ	Bilabial	ɓ	Bilabial	ʼ	as in
ǀ	Dental	ɗ	Dental/alveolar	pʼ	Bilabial
ǃ	(Post)alveolar	ʄ	Palatal	tʼ	Dental/alveolar
ǂ	Palatoalveolar	ɠ	Velar	kʼ	Velar
ǁ	Alveolar lateral	ʛ	Uvular	sʼ	Alveolar fricative

The following are consonants that don't fit into the main IPA chart:

ʍ	Voiceless labial-velar fricative
w	Voiced labial-velar approximant
ɥ	Voiced labial-palatal approximant
ʜ	Voiceless epiglottal fricative[8]
ʢ	Voiced epiglottal fricative
ʡ	Epiglottal plosive
ɕʑ	Alveolo-palatal fricatives
ɺ	Alveolar lateral flap
ɧ	Simultaneous ʃ and x

5. Vowels

Vowels differ from consonants in that they do not have "places of articulation," that is, points of major constriction in the vocal tract. Rather, the vocal tract as a whole acts as a resonating chamber. By modifying the shape of this chamber using movements of the tongue, jaw, and lips, different timbres are imparted to the basic sound produced at the vocal cords.

There are three basic modifications that one can make to the shape of the vocal tract. Vowels are described by specifying each modification used.

5.1 Rounding

One obvious modification one can make to the shape of the vocal tract is to round the lips, thus narrowing the passage at the exit. This happens, for example, in the vowels that many English dialects have for *boot* [u], *book* [ʊ], and *boat* [o]. These are called **rounded** or simply **round** vowels. Other vowels, such as the [i] of *beet* or the [ʌ] of *cut*, are called **unrounded**.

5.2 Height

Another modification one can make to the shape of the vocal tract is to make the passage through the mouth wider or narrower. Widening is accomplished by opening the jaw and/or

[8] Epiglottals are made by touching the upper edge of the epiglottis (see diagram, p. 7) to the rear wall of the pharynx.

lowering the body of the tongue towards the bottom of the mouth. Narrowing is accomplished by raising the jaw and/or raising the body of the tongue.

The terminology for describing these changes is based on the height of the tongue body, without regard to whether this is due to jaw movement or tongue movement. Vowels are classified as **high**, **mid**, or **low**. In effect, high vowels have a narrow passage for the air to pass through, and low vowels have a wide passage. Another terminology, which appears on the IPA chart, is to call the high vowels **close** and the low vowels **open**.

Examples of high vowels in English are [i], the vowel of *b<u>ea</u>t*, and [u], the vowel of (for some English speakers; see below) *b<u>oo</u>t*. Example of low vowels are [ɑ], the vowel of *sp<u>a</u>*, and [æ], the vowel of *b<u>a</u>t*. You can feel the oral passage widening and narrowing if you pronounce a sequence of vowels that alternates between high and low, such as [i æ i æ i æ i æ].

5.3 Backness

The third primary way of changing the vocal tract shape is to place the body of the tongue towards the front part of the mouth or towards the back. Vowels so made are called **front** and **back** vowels, respectively; and vowels that are neither front nor back are called **central**. For example, [i] (*b<u>ea</u>t***)** is a high front unrounded vowel, and the [u] vowel that appears in many languages (e.g. Spanish, French, and Persian) is a high back rounded vowel: French [ʁuʒ] ‘red’.

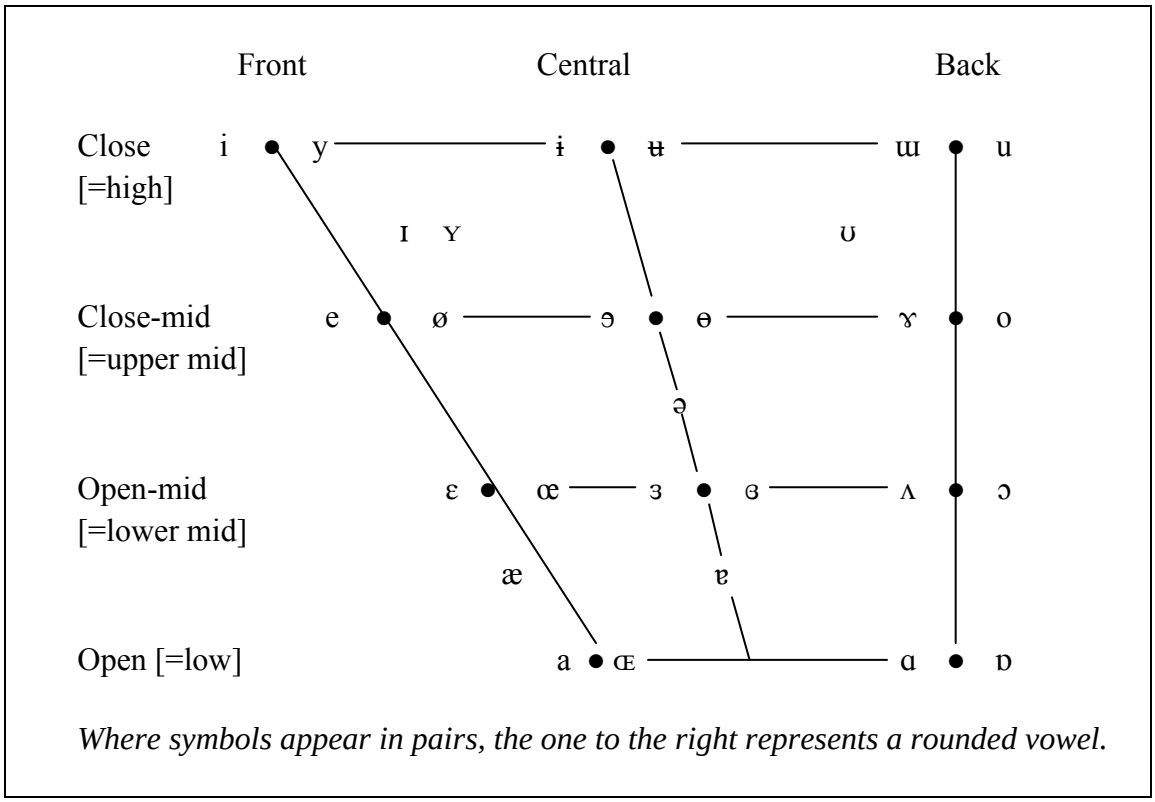
[u] is often described as being the vowel of English words like *b<u>oo</u>t*. This is true, but only for certain dialects of English; other dialects have a vowel that is closer to central than back: [bʉt].

A way to feel backness, particularly if you know how to say a true [u] instead of [ʉ], is to say the sequence [i u i u i u i u …] and feel your tongue body sliding forward and backward along the roof of your mouth.

5.4 Describing vowels systematically

We now have three dimensions for classifying vowels, each based on a particular modification of the vocal tract shape: rounding, height, and backness. The three dimensions allow us to describe vowels clearly and to organize them in a chart.

The IPA chart for vowels looks like this:



Where symbols appear in pairs, the one to the right represents a rounded vowel.

For the use of these symbols in depicting the vowels of English, see p. 24 below.

An awkward problem with the IPA is that there is no symbol for the low central unrounded vowel, which appears to be the most common of all vowels in the world's languages. Below, I will follow the practice of many other linguists in adapting the symbol [a], which strictly speaking designates a front vowel in IPA, to denote the central vowel; where it is crucial, I will state which vowel is being described.

The IPA chart is also a bit puzzling for offering vowel symbols that have no description in terms of categories; for instance [ɪ] floats in the upper left part of the chart without any row or column label. We will remedy this below when we set up a system of **features** (ch. 4) to classify vowels; see p. 87.

5.5 Glide-vowel correspondences

Glides can be described in two ways, because they are essentially the non-syllabic equivalents of vowels (semivowels). Thus, [j] is in IPA terms a palatal central approximant, but it is also describable as a high front unrounded glide, and is thus the consonant counterpart of [i]. Likewise, [w] is a labial-velar central approximant, but it is also treatable as a high back rounded glide, the counterpart of [u].

5.6 Diphthongs

A **diphthong** (note the spelling) is a sequence of two vowels that functions as a single sound. Further, a diphthong always forms just one syllable, whereas a two-vowel sequence forms two. You can transcribe a diphthong by stringing together two vowel symbols denoting its beginning and end, and optionally connecting them with a ligature.

English has numerous diphthongs, the number depending on dialect. The three most noticeable ones are [a͡ɪ], which appears in *ride*; [ɔ͡ɪ], which appears in *boy*; and [a͡ʊ], which appears in *how*. (The diphthong [a͡ʊ] is pronounced [æ͡ʊ] by many speakers.) In addition, most dialects have diphthongs in which the difference between the two component vowels is more subtle, and a rough transcription can use a monophthong symbol. Thus, the sound of *bay*, which some sources transcribe as [e], is [e͡ɪ] for most English speakers, and [o] (*go*) is most often [o͡ʊ]. Other vowels of English can be more or less subtly diphthongal; for example, [ɛ͡ə] for what is normally transcribed [ɛ] (as in *bed*).

5.7 Syllabic consonants

It is possible for sounds that are normally consonants to be prolonged slightly and serve as the nucleus of a syllable. Such sounds are called **syllabic consonants**; they are transcribed by placing the IPA syllabic marker under them, as in [n̩] (*button*, IPA [ˈbʌtn̩]). Generally, only the more sonorous consonants, such as liquids or nasals, can occur as syllabic.[9] Syllabic glides are simply vowels, as noted above in §5.5.

6. Stress and tone

Many languages make distinctions of stress; roughly, the degree of loudness or effort with which a syllable is pronounced. IPA provides the symbol [ˈ] to indicate a strong stress and [ˌ], to indicate a relatively weak (secondary) stress. These marks are placed just before the *syllable*, not the vowel, thus [ˈɹæbɪt] *rabbit*, [əˈba͡ʊt] *about*.

A great many of the world's languages are **tone** languages, using differences of pitch to distinguish words from one another. The IPA offers two distinct systems of tonal transcription. In the chart below, the symbols given first in each category are ordinarily used for languages with mostly level tones; the ones given second are for languages in which the tones often glide up or down within a syllable.

Intonation is use of the voice for linguistic purposes other than distinguishing words; for instance, for distinguish questions from statements. The IPA intonation marks will not suffice for our purposes (see Chapter 15); and are omitted from the following chart.

[9] Various Berber languages have syllabic fricatives and stops.

SUPRASEGMENTALS[10]

			TONES AND WORDS ACCENTS					
			LEVEL			CONTOUR		
ˈ	Primary stress	ˌfoʊnəˈtɪʃən	a̋ or ˥		extra high	ǎ or ˩˥		rising
ˌ	Secondary stress		á	˦	high	â	˥˩	falling
ː	Long	eː	ā	˧	mid	a᷄	˧˥	high rising
ˑ	Half-long	eˑ	à	˨	low	a᷅	˩˧	low rising
˘	Extra-short	ĕ	ȁ	˩	extra low	a᷈	˩˥˩	rising-falling
.	Syllable break	ri.ækt						etc.

7. Diacritics

The IPA provides a number of diacritics that may be attached to symbols to modify their meaning. Of these, the most important ones used here are as follows:

- [ʰ], for **aspiration**. In [pʰ, tʰ, kʰ], aspiration is a puff of breath (which in English can be felt with the fingers, placed in front of the mouth), with a delay in the onset of voicing in a following vowel.
- [◌̥], used to denote **voicelessness** in a symbol that is otherwise interpreted as voiced. For example, [i̥] is a voiceless vowel, found for instance in Japanese.
- [◌˞], the **rhoticity** diacritic, meaning that the tongue tip or blade is curled backward. Rhotacized schwa, [ɚ], is a common way of transcribing the syllabic consonant [ɻ̩]. (Note that this is essentially parallel to the vowel-glide correspondences described in §5.7 above. [ɚ] and [ɻ̩], are essentially the same sound, described from the viewpoint of vowels and consonants.)
- [ː] is placed after vowels (and occasionally consonants[11]) to show they are **long**; likewise [◌̆] is placed over a speech sound to show that it is extra short.
- [◌̃], a tilde placed over a symbol, indicates that it is **nasalized**; that is, pronounced with the velum lowered as in French [mɛ̃] ‘hand’.
- Tiny **adjustments for vowel quality**: [◌˖] a bit higher, [◌˕] a bit lower, [◌˗] a bit backer, [◌˖] a bit fronter. The backer/fronter diacritics may also be used for place of articulation in consonants.

8. Phonetic transcription

The accurate rendering of pronunciation using phonetic symbols is a skill learned through practice; see exercises below. Here I will note only that different transcriptions are appropriate to different purposes.

[10] The two example words in this chart are *phonetician* and *react*.

[11] For consonants, probably the more common practice is to indicate length by doubling.

- On occasion, issues of phonological importance will hinge on matters of tiny phonetic detail. Here, the wide range of diacritics offered by the IPA can be very useful. Transcription that attempts to use symbols to represent speech as accurately as possible is called **narrow** transcription.

- Otherwise, a so-called **broad** transcription, abstracting away from non-crucial detail, usually suffices. Often transcription is narrow in one part of a word (the part containing the matter of interest), but broad elsewhere.

- In selecting a broad transcription, sometimes it is useful to idealize across speakers somewhat. Thus, the various English dialects have a remarkable range of phonetic qualities for the vowel of the word *out* ([a͡ʊ], [æ͡ʊ], [ɑ͡ʊ], [ɑ͡o], [ɛ͡ʊ], [ʌ͡ʊ], etc.), but if one's focus is on other matters, it is harmless to use [a͡ʊ], which is fairly standard in reference sources. The advantage is that the experienced reader can read a standardized transcription more quickly.

9. Exercises

9.1 Web exercises

Phonetic exercises usually involve production or perception, so they are best done on line where sound files can be provided. The exercises for Peter Ladefoged's textbook (see Further Reading, below) are posted at http://www.phonetics.ucla.edu. Some exercises used in the author's own phonetics teaching are available here:

http://www.linguistics.ucla.edu/people/hayes/103/EnglishTranscriptionPractice/
http://www.linguistics.ucla.edu/people/hayes/103/Allophones/
http://www.linguistics.ucla.edu/people/hayes/103/CTranscriptionPractice/

9.2 Study guide questions for this chapter

a. The IPA consonant chart (p. 14) includes a blacked out cell for pharyngeal nasals, claiming that such a sound is "judged impossible." Why is this so? Explain your answer.
b. Same question as (a), but for the voiced counterpart of the glottal stop.
c. Find a pair of contrasting examples showing that we need to be able to transcribe [t͡ʃ] distinct from [tʃ]. Give IPA transcriptions for your examples. (Hint: try stringing words together.)
d. Construct an unambiguous IPA symbol to depict a voiceless dental sibilant affricate, explaining each diacritic that you use.
e. Would it be sensible to use [o̬] in an IPA transcription? Explain your answer.
f. In articulating a velar nasal, the tongue body need not move as far to achieve closure as in a velar stop. Explain why, referring to the figure on p. 7.
g. Give three ways to use the IPA diacritics to transcribe a low central unrounded vowel.
h. Find the errors in the following IPA transcriptions and correct them: *sing* [sɪng], *threat* [thɹɛt], *table* [ˈte͡ɪble], *exit* [ˈɛxɪt], *ballad* [bˈæləd], *heraldry* [ˈhɛɹəldri], *easy* [ízi], *music* [ˈmusɪk].

Further Reading

Some important resources for the International Phonetic Alphabet are:

- The *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet* (1999; Cambridge: Cambridge University Press)
- A Web site containing sound files for all the examples given in the Handbook: http://web.uvic.ca/ling/resources/ipa/handbook.htm.
- The official IPA Web site (http://www2.arts.gla.ac.uk/IPA/ipa.html).
- The widely-used free IPA fonts distributed by the Summer Institute of Linguistics (http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=IPAhome).

An introductory phonetics textbook that I rely on in my own phonetics teaching is *A Course in Phonetics*, by Peter Ladefoged (5th ed. 2005, Heinle). This text is accompanied by downloadable sound files: see http://www.phonetics.ucla.edu. A wide-ranging study of speech sounds is in Peter Ladefoged and Ian Maddieson, *The Sounds of the World's Languages* (1996, Blackwell).

The student who examines journal articles and other reference sources in phonology will encounter, in addition to IPA, a bewildering variety of other phonetic symbols. One also often finds the same symbol used in radically different ways. A useful resource for navigating this thicket is the *Phonetic Symbol Guide*, by Geoffrey Pullum and William Ladusaw (2nd ed. 1996, University of Chicago Press).

A useful Web page on vocal tract anatomy is http://www.phon.ox.ac.uk/~jcoleman/phonation.htm.

Chapter 2: Phonemic Analysis

1. Phonology and phonetics

As noted in the previous chapter, there are two branches of linguistics that deal with speech sounds. Phonetics studies speech sounds in ways that are close to the speech stream, focusing on production, acoustics, and perception. Phonology tends to be more abstract, dealing not directly with the physical nature of speech sounds (though that is of course quite relevant), but rather with the largely unconscious **rules** for sound patterning that are found in the mind/brain of a person who speaks a particular language. It could be said that a phonologist is a kind of grammarian, and the area of grammar that she studies is the sound pattern of a language.

The rules studied by phonologists come in various kinds. First, phonetic study shows that sounds *vary with their context*, often in complex ways; and phonologists hypothesize rules to characterize this variation. Second, the *sequencing and distribution* of speech sounds is not arbitrary, but follows patterns also describable with rules. Third, phonology is *interfaced* with other components of the grammar, particularly morphology and syntax, and there are rules that characterize the way in which sound patterning reflects information that arises within these components.

The phonologies of many languages often show a level of complexity that make them a worthwhile intellectual challenge for the phonologist trying to understand them. It can take many years of careful research to fully explicate the sound pattern of a language. What is remarkable is that the same pattern is learned quite rapidly, at the intuitive level, by humans when they are exposed to it in childhood.

2. Distinctiveness and contrast

The sounds of a language are intrinsically meaningless: their only purpose is to form the building blocks of which words are made. For example, because English has the sounds [t] and [d], the possibility exists of English having the word *time* [taɪm], distinct from the word *dime* [daɪm]. One could put it this way: the only real purpose of a speech sound is to sound different from the other sounds of the language; this is what makes a spoken vocabulary possible.

To begin the analysis of a language's phonology, we therefore seek to locate all of its basic sounds, the minimal units that serve to distinguish words from each other. These basic speech sounds are the **phonemes** of the language. The phonemes of one commonly-spoken dialect of American English are arranged phonetically below; that is, more or less in the manner of the IPA chart, though not necessarily the exact same rows, columns, or order that the IPA uses. The sound symbols are in slant brackets, which is the standard way of indicating phonemes. Example words are given beneath each phoneme to illustrate it.

Consonants

		Bilabial	*Labio-dental*	*Dental*	*Alveolar*	*Palato-alveolar*	*Palatal*	*Velar*	*Glottal*
Stops	*voiceless*	/p/ p<u>i</u>n			/t/ t<u>i</u>n			/k/ k<u>i</u>n	
	voiced	/b/ b<u>i</u>n			/d/ d<u>i</u>n			/g/ g<u>i</u>ll	
Affricates	*voiceless*					/tʃ/ ch<u>i</u>n			
	voiced					/dʒ/ g<u>i</u>n			
Fricatives	*voiceless*		/f/ f<u>i</u>n	/θ/ th<u>i</u>n	/s/ s<u>i</u>n	/ʃ/ sh<u>i</u>n			/h/ h<u>y</u>mn
	voiced		/v/ v<u>i</u>m	/ð/ th<u>i</u>s	/z/ z<u>i</u>p	/ʒ/ vi<u>s</u>ion			
Nasals		/m/ m<u>i</u>tt			/n/ n<u>i</u>p			/ŋ/ si<u>ng</u>	
Approxi-mants	*lateral*				/l/ <u>L</u>ynn				
	central	/w/ w<u>i</u>n			/ɹ/ r<u>i</u>m		/j/ y<u>i</u>ng		

Vowels and Diphthongs

	Front	*Central*	*Back*		*Diphthongs*
	Unrounded	*Unrounded*	*Unrounded*	*Rounded*	
Upper high	/i/ b<u>ea</u>t			/u/ b<u>oo</u>t	/aɪ/, /aʊ/, /ɔɪ/ b<u>i</u>te, b<u>ou</u>t, C<u>oi</u>t
Lower high	/ɪ/ b<u>i</u>t			/ʊ/ f<u>oo</u>t	
Upper mid	/eɪ/ b<u>ai</u>t	/ə/ abb<u>o</u>t		/oʊ/ b<u>oa</u>t	*Rhotacized upper mid central unrounded* /ɚ/ B<u>er</u>t
Lower mid	/ɛ/ b<u>e</u>t		/ʌ/ but	/ɔ/ b<u>ough</u>t	
Low	/æ/ b<u>a</u>t		/ɑ/ f<u>a</u>ther		

Other English dialects differ from the above, having additional phonemes such as /ʍ/, /ʌɪ/, /ɛə/, /ɒ/; or fewer phonemes.

Languages vary in their number of phonemes. The record low is believed to be held by Rotokas (East Papuan, New Guinea), with 11, and the record high by !Xóõ (Khoisan,

Botswana/Namibia), with 160. English has roughly 37-41, depending on the dialect and the analysis. The average across languages is about 30.

Below, we will discuss detailed methods for establishing the phoneme inventory of a language or dialect. But the most important point can be stated right away: *if any two words of a language are pronounced differently, they must differ in at least one phoneme*. This follows from the basic idea of the phoneme, that is, that the phoneme inventory is the set of "building blocks" out of which all the words of the language are constructed.

The example given above, *time* [taɪm] vs. *dime* [daɪm], was strategically arranged to make this point. These words are identical, except for their initial sounds; that is, they are both of the form [Xaɪm]. Since they are different words, it follows that [t] and [d] are distinct sounds; that is, they are separate phonemes. A pair like ([taɪm], [daɪm]), differing in just one single location, is called a **minimal pair**. A minimal pair is the most effective way to show that two sounds are distinct phonemes.

There are quite a few ways in phonology of saying that two sounds are separate phonemes. Equivalently, we say that the English sounds /t/ and /d/ **contrast** with each other; that they are **in contrast**; or that they are **phonemically distinct**; or that the difference between them is **distinctive**. All of these terms are essentially equivalent.

The concept of minimal pair can be extended to cover larger sets. A set like *time* [taɪm] – *dime* [daɪm] – *lime* [laɪm] is a **minimal triplet**, showing that /t/, /d/, and /l/ are distinct phonemes; and the concept clearly generalizes to as many members as one can find. The consonant chart above includes examples forming a minimal 13-tuplet for consonants and a minimal 12-tuplet for vowels. Such sets are useful for demonstrating a large fraction of the phonemic system of a language all at once.

3. Sounds that do not contrast

For a reason to be given, there are also many pairs of sounds (in any language) that do *not* contrast. Here is a simple case from English, involving the length of vowels. If you listen to a native speaker say the following pairs of words (or better, measure with acoustic equipment), you will find that the vowel phoneme /eɪ/ is quite a bit **shorter** in the second member of each pair. I've indicated this in the transcription with the IPA shortness marker on the [e] part of the [eɪ] diphthong:

save	[seɪv]	*safe*	[sĕɪf]
Abe	[eɪb]	*ape*	[ĕɪp]
made	[meɪd]	*mate*	[mĕɪt]
maze	[meɪz]	*mace*	[mĕɪs]
age	[eɪdʒ]	*H*	[ĕɪtʃ]
Haig	[heɪg]	*ache*	[ĕɪk]

Although [eɪ] and [ĕɪ] are audibly different, they are not separate phonemes—one could not use them to form a distinction between words. The reason is that their distribution is predictable. In the data given, which are representative, there is a straightforward fact that determines which of the two will appear. (You should take a look at the data now if you have not yet seen what this factor is.)

The relevant factor is the voicing of the immediately following sound. [eɪ] occurs when this sound in the word is voiced (here: [v, b, d, z, dʒ, g]), and [ĕɪ] occurs when the next sound in the word in voiceless (here: [f, p, t, s, tʃ, k]). The fact that the appearance of [eɪ] vs. [ĕɪ] is predictable is important, because it shows that the difference between the two could never be the (sole) distinction between words; there will always be a difference in the voicing of the following consonant as well. It follows that there can be no minimal pair for [eɪ] and [ĕɪ].

A term that is commonly used to describe this is **complementary distribution**: two sounds are said to be in complementary distribution if one sound never occurs in the environments in which the other occurs.

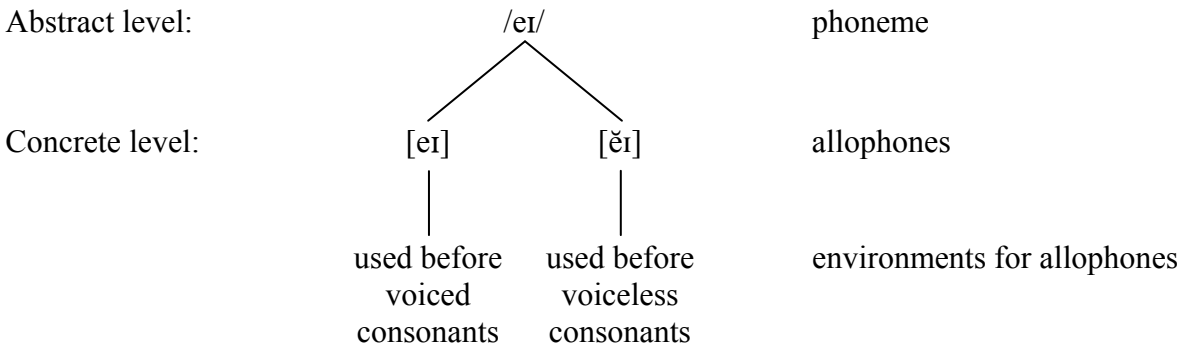
Thus, in phonological analysis, for any pair of sounds it is necessary to establish their phonological status: either they are separate phonemes, capable of distinguishing words, or mere variants, whose distribution in the language is determined by context, in a way that can be expressed by a rule (here, the rule relating length to voicing). We will see refinements on this point later on, but it will suffice for now.

To complete the description of [eɪ] and [ĕɪ], we must dispose of an alternative possibility: that [eɪ] and [ĕɪ] really are distinct phonemes, and it is the voicing of the following consonant that is predictable. This possibility is eliminated by the fact that minimal pairs occur for consonant voicing in other contexts (for example, *few* vs. *view*); thus it has to be the voicing that is phonemic and the length that is predictable.

4. Phonemes as categories

Another important aspect of the [eɪ]-[ĕɪ] data under discussion is that virtually every English speaker is unaware of the difference until it has been pointed out. That is to say, speakers are

willing, intuitively, to accept [eɪ] and [ĕɪ] as being the "same vowel".[12] Phonologists hypothesize that sounds [eɪ] and [ĕɪ] in the present case (and similarly in parallel cases) form an abstract phonological *category*, namely, the phoneme /eɪ/. The concrete, observable sounds [eɪ] and [ĕɪ] are called the **allophones** of /eɪ/. This is illustrated as follows.



The idea is that the fundamental phonological categories (the phonemes) can be used to distinguish words from each other, but the variants of a particular phoneme (the allophones), cannot. As a metaphor, you could imagine that the phoneme inventory of a language is the fundamental "alphabet" (an alphabet of sound) out of which all the words of a language are composed; but each letter is subject to contextual variation. At the level of conscious awareness, people are characteristically attuned only to the distinctions between phonemes; to make people aware of allophones requires that their attention be carefully directed to the distinction.

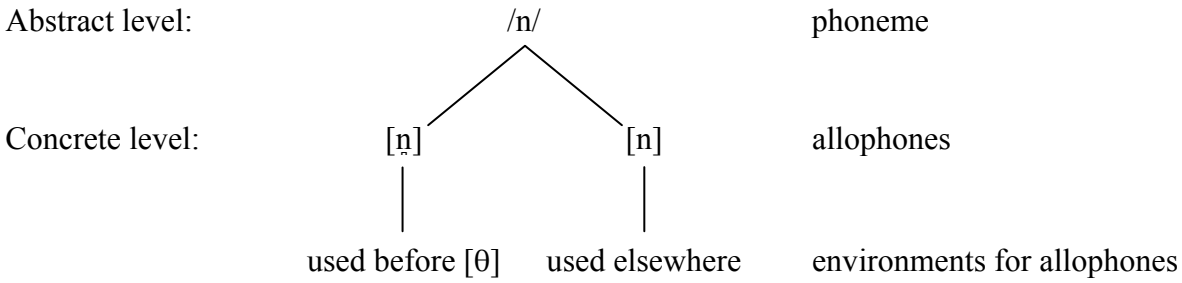
5. More instances of allophonic variation

Before moving on, let us consider some other cases of allophonic variation in English. The following pair illustrates words containing alveolar [n] and dental [n̪]. Check the environments for each sound, establishing the complementary distribution, before you read further.

Words with [n]		**Words with [n̪]**	
know	[ˈnoʊ]	*tenth*	[ˈtɛn̪θ]
annoy	[əˈnɔɪ]	*month*	[ˈmʌn̪θ]
onion	[ˈʌnjən]	*panther*	[ˈpæn̪θɚ]
nun	[ˈnʌn]	*chrysanthemum*	[kɹəˈsæn̪θəməm]

It is not hard to see that the dental [n̪] occurs in a specific context: before [θ]. There is no particular context for alveolar [n]; it occurs pretty much everywhere else. Thus, the phonemic pattern is as follows

[12] This is true even when the sounds are spelled differently, as in *Haig* vs. *ache*. The intuitive judgments are of sound, not spelling.

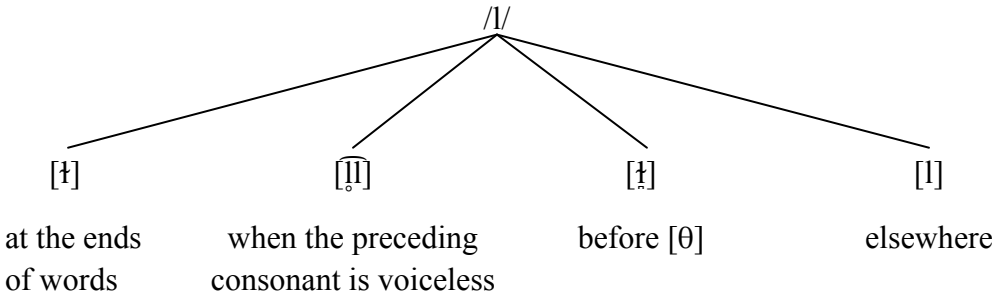


The "elsewhere" environment seen here is quite common in phonology, and cases like the [n] in this example are often called **elsewhere allophones**. The allophone [eɪ], seen in the previous example, is actually an elsewhere allophone; it occurs not just before voiced consonants, but at the end of a word, as in *bay* [beɪ] or *day* [deɪ].

The next data set illustrates four allophones of the /l/ phoneme as they occur in a number of dialects of English. [ɬ] is a velarized l, articulated with high back tongue body position. [ɬ̪] is the same as [ɬ], only with a dental instead of alveolar place of articulation. [l̥͡l] is an l which starts out voiceless and ends voiced. Before you read further, inspect the following data and determine the environment characterizing each sound.

Words with [ɬ]		**Words with [l̥͡l]**		**Words with [ɬ̪]**		**Words with [l]**	
file	[ˈfaɪɬ]	*slight*	[ˈsl̥͡laɪt]	*wealth*	[ˈwɛɬ̪θ]	*listen*	[ˈlɪsən]
fool	[ˈfuɬ]	*flight*	[ˈfl̥͡laɪt]	*health*	[ˈhɛɬ̪θ]	*lose*	[ˈluz]
all	[ˈɔɬ]	*plow*	[ˈpl̥͡laʊ]	*filthy*	[ˈfɪɬ̪θi]	*allow*	[əˈlaʊ]
ball	[ˈbɔɬ]	*cling*	[ˈkl̥͡lɪŋ]	*tilth*	[ˈtɪɬ̪θ]	*aglow*	[əˈgloʊ]
fell	[ˈfɛɬ]	*discipline*	[ˈdɪsəpl̥͡lən]	*stealth*	[ˈstɛɬ̪θ]	*blend*	[ˈblɛnd]
feel	[ˈfiɬ]						

The pattern can be described as follows.



As before, this description appears to hold for the entire language, not just the sample data given here. Since none of these environments overlap,[13] the description establishes complementary distribution, and we can claim that all four of these sounds are allophones of the same phoneme.

The three examples just given are just the tip of the iceberg; in fact, virtually all the phonemes in English show variation based on their context. If we looked at English in full phonetic detail, taking all the allophonic variation into account, we would find that it has not several dozen speech sounds, but thousands.

All this gives rise to an overview of how a phonological system is "designed". In every language, the number of sounds that can be uttered is very large. But the phonological system organizes these sounds in a particular way, such that only a small subset of phonetic differences (for example, in English [t] vs. [d], or [ɪ] vs. [ɛ]) can serve to distinguish words. The remaining phonetic differences are allophonic, and regulated by rule.

6. Phonemic transcription

When a linguist records words as sequences of phonemes (under a particular phonemic analysis), the result is termed a **phonemic transcription**. This is to be distinguished from a **phonetic transcription**, which includes allophonic detail—the degree of detail recorded being up to the transcriber. The term **orthographic transcription** simply means that the words are written down using the customary spelling system (orthography) of the language. Below I give the same sentence in orthographic, phonemic, and phonetic transcription.

This is an orthographic transcription.
/ðɪs ɪz ə foʊˈnimɪk trænˈskɹɪpʃən/ (This is a phonemic transcription.)
[ˈd̪ðɪs ɨz ə fə̃ˈnɛɾɨˀk t͡sʰɹɛə̃nˈskɹɪpʃɨ̃n] (This is a phonetic transcription.)

The attractiveness of a phonemic transcription for practical purposes is that it is far simpler than a phonetic transcription, yet (provided one knows the rules) it conveys the same information. One need only apply the rules to derive the correct allophones.

Reference grammars (books addressed to linguists that offer a description of a language) often begin by setting out the phonemes and allophones. The first few transcriptions in a grammar are usually phonetic; then, once the allophonic rules have been duly set out, all the remaining transcriptions can safely be phonemic, without any loss of information.[14]

7. Phonological rules

Generalizations about the patterning of allophones can be stated as phonological **rules**. For instance, to describe the patterning of [eɪ] and [ĕɪ] given above, one might write a rule like this:

[13] The environments "after a voiceless consonant" and "at the end of a word" cannot overlap, because English has no words ending in a voiceless consonant followed by /l/.

[14] For convenience, authors of reference grammars usually take the further step of setting up a **practical orthography**, in which each phoneme is spelled using an ordinary letter or letter combination.

/eɪ/ Shortening

The phoneme /eɪ/ *is realized as extra short when a voiceless consonant follows.*

We will refine our rules in many ways below, but this should get across the basic idea. The concept of rule is central to phonology; here are some elaborations.

First, rules are **language-specific**: the shortening of /eɪ/ (and, as it turns out, of other vowels) must be considered as a rule *of English*; it is not a universal rule, nor some kind of general principle of speech articulation. We know this because we have data from other languages that apparently lack any rule of this kind. For instance, neither Polish nor Saudi Arabic shortens vowels before voiceless consonants. The shortening rule of English is part of the phonological pattern of the English language, and must be learned in some form by children acquiring English.

Second, rules are usually **productive** in the sense that they extend to novel cases. "*Vake*" and "*praig*" are not words of English, but if they become words, we can be confident that they would obey the rules and be pronounced [vĕɪk] and [pɹeɪg]. For more on productivity, see Chapter 9.

Third, rules give rise to **well-formedness intuitions**. If a phonetician, or a speech synthesizer, were to create exceptions to the rule, English speakers sense the awkwardness of the result; thus [sĕɪv] and [seɪf] are inappropriate as natural renditions of *save* and *safe*. In other words, rule violations are sensed intuitively.

Fourth, phonological rules are **untaught**. Instead, they are learned intuitively by children from the ambient language data, using mechanisms that are as yet unknown. In this respect, phonological rules are very different from rules that are imparted by direct instruction, like (for example) the rules for traffic lights, or rules of **normative grammar** like "don't end a sentence with a preposition."

Lastly, phonological rules are evidently a form of **unconscious knowledge**. No matter how hard we try, we cannot access our phonological rules through introspection.

One shouldn't be surprised that this is so, because most of the computations that our brains carry out are similarly inaccessible to consciousness. For example, we can detect color constancy under variable conditions of light and shadow, or the direction of sound sources by the time delay between our ears. These mental processes involve rapid, automatic mental computations that cannot be intuited by the conscious mind as they occur. We consciously notice the *result* of such computations ("this object is uniformly red"; "a car is approaching from my left"), but not the way it is done. To understand such processes, cognitive scientists *infer* their mechanisms on the basis of observation, experimentation, and theorizing. No one bothers to ask people how they do these things, because people don't know.

Phonology is similar. When we speak, we automatically obey hundreds, perhaps thousands of phonological rules, but we can neither observe nor articulate what these rules are. Thus, when this book discusses "rules", what is meant is rules of the unconscious kind. We cannot learn about these processes through introspection, but must proceed indirectly, through data gathering, experiment, and construction of theories.

8. Formalizing phonological rules

We turn now to the problem of expressing the phonological rules precisely. In principle, we could just write all of the rules in prose—and indeed, this is usually done as backup, to help the reader understand the rules more easily. But in general, phonologists have found that use of a formal notation permits greater precision and clarity. Throughout this book, we will gradually accumulate more notational apparatus with this purpose in mind. The notations used here are drawn from the research literature in phonology; I have tried to limit myself to notations that would be widely recognized among phonologists.

8.1 Expressing environments

Let us start with formalism for describing the environments where allophones occur. The symbol **slash**, "/", as used in phonology, means "in the environment." A **long underline** stands for where the allophone occurs relative to its neighbors. Thus the following expression:

/ ___ θ

is to be read "in the environment 'before theta'", or for short, just "before theta." If instead we had written "/ θ ___," it would be read "after theta".

In expressing the environment of an allophone, we often must specify not just a single sound like [θ], but a whole class of sounds. For example, the environment for [l͡l̥] (p. 28) includes the class of voiceless consonants. To describe such classes, we use **square brackets**, containing the particular phonetic properties—which, in the context of phonology, are called **features**—that designate the relevant class of sounds. (Features are covered in detail in ch. 4.) Thus, the following notation can be read "after a voiceless consonant":

$$/\begin{bmatrix}\text{consonant}\\ \text{voiceless}\end{bmatrix}\text{\_\_\_}$$

As can be seen, square brackets in phonology essentially mean "*and*"; hence $\begin{bmatrix}\text{consonant}\\ \text{voiceless}\end{bmatrix}$ means "a segment[15] which is a consonant *and* is voiceless."

The symbols "+" and "–" are used before feature names to mean that a segment either has, or does not have, the phonetic property that a feature designates. Thus, in more standard notation the environment just given would appear as:

[15] As phonologists generally do, I will use the term "segment" to refer to a single speech sound.

$$/ \begin{bmatrix} +\text{consonant} \\ -\text{voice} \end{bmatrix} \_\_$$

Where we want to refer to the beginnings and ends of grammatical constituents like words, we can use brackets, much as is done in the study of syntax and morphology. For example, the notation given below can be read "at the end of a word."

$$/ \_\_\_]_{\text{word}}$$

"At the beginning of a word" would be / $[_{\text{word}}$ ___.

8.2 Underlying representations and derivations

We turn next to the task of characterizing allophones as the variants of a single abstract phoneme. A widely adopted theoretical approach in phonology is to characterize the phoneme by setting up an abstract level of representation called the **underlying representation**, also called the **phonemic representation**, **underlying form** or **base form**. The idea is that phonemes have an essential, characteristic form, which is altered in particular contexts by the rules of the phonology, applying in a derivation.

In a system of this kind, it is rational to adopt as the underlying representation of the phoneme its "elsewhere" allophone. Recall (p. 28) that the elsewhere allophone is the allophone that is not affiliated with any particular context, but rather is the sound that appears when no other special context is met. The phonological derivation starts out with the underlying form, and rules apply to derive from it the various allophones in their appropriate contexts. If no rule is applicable, the underlying form emerges unaltered as the output of the phonology.[16]

Using this approach, we can develop an explicit description of the system of allophones for the English phoneme /l/. We select /l/ as the underlying representation, and posit three rules, stated below in both formalism and prose.

/l/ Devoicing

$$/l/ \rightarrow [\widehat{l\underset{\circ}{l}}] \; / \begin{bmatrix} +\text{consonant} \\ -\text{voice} \end{bmatrix} \_\_$$

Partially devoice /l/ *after a voiceless consonant*.

[16] A caution: I find that students sometimes spontaneously adopt a terminology in which the elsewhere allophone is termed "the phoneme" and the contextual allophones derived by rule are called the "allophones". This is perfectly coherent, but is not standard usage. Among phonologists, the elsewhere allophone counts as an allophone just like all the others, and the phoneme is a separate, abstract entity—it occurs at a deeper level of representation.

/l/ Dentalization

/l/ → [l̪] / ___ θ

/l/ *is rendered as velarized and dental before* [θ]

/l/ Velarization

/l/ → [ɬ] / ___]$_{word}$

/l/ *is velarized word-finally*.

Along with the posited underlying forms and rules, an analysis of this type is usually illustrated by providing sample **derivations**. A derivation consists of a series of lines. The first contains the underlying representations of a set of forms, and the last contains the actual phonetic forms, which in this context are often called **surface representations**. The intermediate lines show the application of the rules in order. Where a rule is inapplicable, the notation "—" is used to designate this.

Here is a derivation for four words containing /l/, specifically chosen to illustrate all of the rules above.

file	*slight*	*wealth*	*listen*	
/faɪl/	/slaɪt/	/wɛlθ/	/ˈlɪsən/	underlying forms
—	s͡l̥laɪt	—	—	/l/ Devoicing
—	—	wɛɬ̪θ	—	/l/ Dentalization
faɪɬ	—	—	—	/l/ Velarization
[ˈfaɪɬ]	[s͡l̥laɪt]	[ˈwɛɬ̪θ]	[ˈlɪsən]	surface forms

In this approach, we need not specify that the elsewhere allophone is [l]; that is simply the base form whenever none of the phonological rules happen to alter it. In other words, a phonological rule like "/l/ → [l] / elsewhere" is unnecessary.

The idea of a phonological derivation has over time proven fruitful. Often, the rules apply in an intricate, cross-cutting pattern, creating large numbers of allophones.[17] It also turns out (Chapter. 7) that in many cases, the order in which the rules apply is crucial.

The derivations form the heart of a phonological discussion, and the reader of a phonological analysis is well advised to inspect rather than skim them.[18] In particular, in each

[17] For instance, we can note that /l/ Dentalization probably doesn't need to carry out the full change /l/ → [l̪]; rather, it should only make the change /l/ → [l̪], and a suitably generalized version of /l/ Velarization can handle that part of the change that velarizes the /l/.

[18] I find that in hard cases it is helpful to copy them down.

case it is important to understand, by comparing the rule to the form, why the rule applied or did not apply. Thus, in reading the first line of the derivation above, you would want to reassure yourself that /l/ Devoicing did indeed apply correctly to *slight*, because /s/ is a member of the class of voiceless consonants; and similarly in all other cases.

9. Phonemes in other languages

A great number of languages have been subjected to phonemic analysis. This typological study has found great diversity, but also a certain degree of unity. As an example of the latter, there is a certain "core" set of speech sounds that tend to be employed as phonemes in a great number of languages. The following set, for example, constitutes all the sounds that occurred in at least 40% of the languages in Maddieson's (1984) survey of phonemic systems. It might be thought of as a "maximally ordinary" phonemic system.

p	t	tʃ	k		i	u
b	d	dʒ[19]	g		e	o
f	s	ʃ			a	
m	n		ŋ			
		l				
		ɾ				
w		j				

On the other hand, most phonemic inventories are not restricted to just these "core" sounds; more normally, an inventory will contain additional, more unusual sounds. For example, the typologically-unusual aspects of English include /θ/, /ð/, the syllabic consonants, and the heavy representation of diphthongs in the vowel inventory. Unusual sounds often occur in multiple languages in the same geographic area; e.g., retroflexes in India and Australia, diphthongs in Northern Europe, and gliding tones in East Asia.

However, phonemic diversity extends beyond just phoneme inventories. A more subtle cross-linguistic difference concerns how the phonetic inventory of a language (that is, the complete collection of allophones) is organized into phonemes. In particular, a distinction that is phonemic (serves to distinguish words) in one language might be allophonic (predictably distributed) in another.

An example is found in the phonemic systems of English and Spanish. Spanish has many sounds that resemble sounds of English (we will consider only North American dialects of English here). In particular, English has a [t] and a tap [ɾ]. The [d] of Spanish is dental rather than alveolar, and there are also slight differences in the tap, but these are small enough to ignore for our purposes.

[19] [dʒ] actually falls somewhat short of 40%; it is included in the list above because of another strong crosslinguistic tendency; i.e. for the sounds to occur in complete, symmetrical series (e.g., voiced matching up with voiceless).

In North American English [ɾ] is (to a rough approximation) an allophone of the /t/ phoneme. The environment for [ɾ] is between two vowels of which the second is stressless. Words having /t/ that fit this environment, and which therefore show a tap, are given in the first column below.

	Phonemic	**Phonetic**		**Phonemic**	**Phonetic**
data	/ˈdeɪ**t**ə /	[ˈdeɪ**ɾ**ə]	*tan*	/ˈ**t**æn/	[ˈ**t**æn]
latter	/ˈlæ**t**ɚ/	[ˈlæ**ɾ**ɚ]	*attend*	/əˈ**t**ɛnd/	[əˈ**t**ɛnd]
eating	/ˈi**t**ɪŋ /	[ˈi**ɾ**ɪŋ]	*guilty*	/ˈgɪl**t**i/	[ˈgɪl**t**i]
Ottoman	/ˈɑ**t**əmən/	[ˈɑ**ɾ**əmən]	*cat*	/ˈkæ**t**/	[ˈkæ**t**]
rhetoric	/ˈɹɛ**t**əɹɪk/	[ˈɹɛ**ɾ**əɹɪk]	*active*	/ˈæk**t**ɪv/	[ˈæk**t**ɪv]
automatic	/ˌɔ**t**əˈmæ**t**ɪk/	[ˌɔ**ɾ**əˈmæ**ɾ**ɪk]	*Atkins*	/ˈæ**t**kɪnz/	[ˈæ**t**kɪnz]

The second column combines other allophones of /t/, without narrowly transcribing their specific properties. In this column, we see where /t/ does **not** appear as the [ɾ] allophone: either because it fails to follow a syllabic sound (*tan*, *guilty*, *active*) or because it fails to precede a syllabic sound (*cat*, *Atkins*), or because the following syllabic sound is stressed (*attend*). But if all the right conditions are met simultaneously, as in the first column, we get [ɾ].

It can be seen that the difference between [t] and [ɾ] is not distinctive in English: the tap is a conditioned variant of the /t/ phoneme that shows up in a particular environment. The Tapping rule can be stated, as a first approximation, as follows:

Tapping

$$/t/ \rightarrow [ɾ] \,/\, [+\text{vowel}] \_\_ \begin{bmatrix} +\text{vowel} \\ -\text{stress} \end{bmatrix}$$

The phoneme /t/ *is realized as* [ɾ] *when it is preceded by a vowel and followed by a stressless vowel.*

Here are derivations:

data	*tan*	*attend*	*cat*	*guilty*	
/ˈdeɪ**t**ə /	/ˈ**t**æn/	/əˈ**t**ɛnd/	/ˈkæ**t**/	/ˈgɪl**t**i/	underlying forms
ɾ	—	—	—	—	Tapping
[ˈdeɪɾə]	[ˈ**t**æn]	[əˈ**t**ɛnd]	[ˈkæ**t**]	[ˈgɪl**t**i]	surface forms

In Spanish, /t/ and /ɾ/ are separate phonemes, as is demonstrated by minimal pairs such as the following:

[ˈpita] ‘century plant’
[ˈpiɾa] ‘funeral pyre’

As with the minimal pairs given for English above, this one demonstrates that for Spanish, the difference between [t] and [ɾ] signals a difference in meaning. That is to say, [t] and [ɾ] are in contrast, and are separate phonemes, /t/ vs. /ɾ/.

Comparing English and Spanish, we see that the [t] vs. [ɾ] difference is allophonic (non-distinctive) for English, but phonemic (distinctive) for Spanish. Thus, in this area, the two languages are *phonetically similar but phonologically different*.

Here is a similar case. Both English and Spanish have a [d] and a [ð] (the voiced dental fricative). In English, we know that the two sounds are separate phonemes, because minimal pairs exist:

die	[**d**aɪ]	vs.	*thy*	[**ð**aɪ]
bayed	[beɪ**d**]	vs.	*bathe*	[beɪ**ð**]
den	[**d**ɛn]	vs.	*then*	[**ð**ɛn]

But in Spanish, there are no such pairs. Furthermore, by looking at Spanish data one can determine that [d] and [ð] are allophonic variants:

[daðo]	‘given’
[deðo]	‘finger’
[usteð]	‘you (polite)’
[donde]	‘where’
[de ðonde]	‘from where’

These and other data indicate that [ð] occurs only after a vowel, while [d] is the elsewhere allophone, occurring after consonants and initially. Thus [ð] and [d] are allophones of the same phoneme.

We can set up the following phonological analysis for the sounds of Spanish discussed so far.

Phonemes: /t/, /d/, /ɾ/

Phonological rule: **/d/ Spirantization**

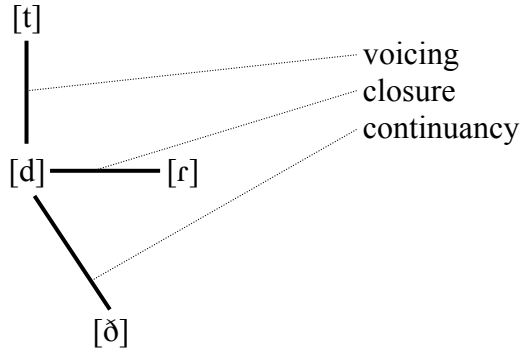
/d/ → [ð] / [+vowel] ___

The phoneme /d/ *is realized as* [ð] *when it follows a vowel*.

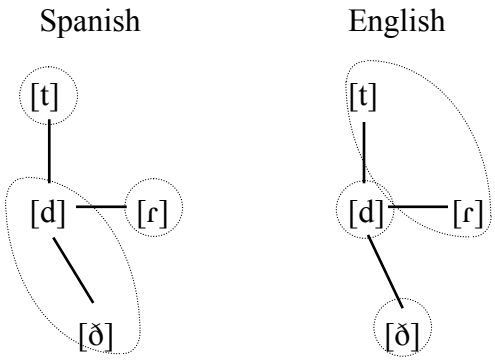
Derivations:	'given'	'you'	'where'	
	/dado/	/usted/	/donde/	underlying forms
	ð	ð	—	/d/ Spirantization
	[daðo]	[usteð]	[donde]	surface forms

Regarding the name of the rule, **spirantization** is the conventional term in phonology for rules that convert stops to fricatives; such rules are common. "Spirant" is a mostly-obsolete synonym for "fricative".

The differences in phonological organization between English and Spanish reflect a different division of phonetic space. Suppose we construe phonetic space as made up of multiple dimensions. We place [d] at the center of this space, and in different directions show [ð] as differing from [d] minimally in its fricative character ("continuancy"); [ɾ] differing from [d] in having short, weak closure; and [t] differing minimally from [d] in voicing:



The phones of this phonetic space are grouped into phonemes differently by Spanish and English, as shown below:



The dotted lines surround groups of sounds that fail to contrast, and thus form single phonemes in the language in question. English has /d/, /ð/, and /t/, with the latter having two allophones [t] and [ɾ]. Spanish has /t/, /ɾ/, and /d/, with the latter having two allophones [d] and [ð]. The chart shows that the sound systems of languages can differ in their phonological organization, as well as in the sounds that they contain. In principle, we could imagine two languages that had exactly the same sounds, but a radically different phonological organization. This would happen if the two languages selected different phonetic distinctions to be contrastive vs. non-contrastive. Using the phonemic method, we would analyze two such languages as having the same set of sounds, grouped into phonemes in two different ways.

10. Phonemicization

Phonemicization is the body of knowledge and techniques that can be used to work out the phonemic system of a language. The method described below has been in existence for several decades and has been used on many languages. Of course, no recipe in linguistics provides certain results, and later on in this text we will see cases where the method falls short. But it is usually the starting point for working out the phonology of a language.

A really solid phonemicization is often the result of years of hard work, carried out by linguists with good ears and extensive experience with the target language. The reason that phonemicization takes so long is that the first linguist or team of linguists to encounter a language will quite often fail to notice a difficult-to-hear contrast. Another factor is that certain phonemes might be rare, and will be encountered only after the linguist has collected a large vocabulary.

10.1 Minimal pairs

By far the most effective method in phonemicization is to look for **minimal pairs**, which (to review) are defined as two different words that differ in exactly one sound in the same location. Some examples: *sip* [sɪp] and *zip* [zɪp] in English form a minimal pair for the phonemes /s/ and /z/; *sill* [sɪl] and *zeal* [zil] are not a minimal pair, because they differ in two locations; *seal* [sil] and *eels* [ilz] are not a minimal pair because the /s/ and /z/ occur in different places. Two sounds that appear in a minimal pair are almost always distinct phonemes.[20]

[20] The exceptions are discussed below in ch. 7, §1.1, and ch. 10, §3.

The **absence** of a minimal pair does not prove much. Often, a language will lack minimal pairs for a pair of relatively rare phonemes simply by accident. A method for dealing with such cases is given in the next section.

As noted earlier, minimal pairs generalize to minimal triplets, quadruplets, and so on. Often, selecting a good 'frame' or phonological context will make it possible to justify quite a bit of the phonemic inventory of a language. A notation that is commonly used for such frames is to place the environment sounds on either side of an underlined blank, which represents the sound being manipulated in the pair, quadruplet, etc. Thus for American English vowels, the frame / h __ d/ gets all but /ɔɪ/, though admittedly some of the words are a bit forced: [21]

heed	[hid]			*who'd*	[hud]
hid	[hɪd]			*hood*	[hʊd]
hayed	[heɪd]			*hoed*	[hoʊd]
head	[hɛd]	*HUD*	[hʌd]		
had	[hæd]	*hod*	[hɑd]	*hawed*	[hɔd]
hide	[haɪd]	*how'd*	[haʊd]		
heard	[hɚd]				

For the missing /ɔɪ/, it is easy to imagine that "*hoid*" **could** be a word; its absence from English is essentially an accident.

10.2 Near-minimal pairs

There are cases in which it is impossible to find minimal pairs for a phoneme. This probably occurs more frequently in languages with long words and large phoneme inventories. In English there appear to be cases where, at least for some idiolects, a minimal pair cannot be found. [22] Conducting a search in an electronic dictionary for minimal pairs for English /ð/ and /ʒ/, I found that it included only three plausible candidates:

bathe	[ˈbeɪð]	vs.	*beige*	[ˈbeɪʒ]
leather	[ˈlɛðɚ]	vs.	*leisure*	[ˈlɛʒɚ]
seethe	[ˈsið]	vs.	*siege*	[ˈsiʒ]

[21] *hayed* 'made hay', *HUD* 'colloquial abbreviation for United States Department of Housing and Urban Development'

[22] An **idiolect** is a language as it is learned and internalized by a single individual. A dialect is a collection of closely similar idiolects, characterizable by region or social class. I refer to idiolects here because I doubt that there is any English dialect whose speakers all happen to lack minimal pairs of the kind under discussion.

However, for all three [ʒ] words, the pronunciation varies by dialect: there are many speakers who have [ˈbeɪd͡ʒ] for *beige*, [ˈliʒɚ] or [ˈlɛʒʊɹ] for *leisure*, and [ˈsid͡ʒ] for *siege*. For a speaker who employs these pronunciations for all three words, there are presumably no minimal pairs for [ð] vs. [ʒ].

Despite this, it is impossible that [ð] and [ʒ] could be allophones of the same phoneme, even in such a dialect. If they were allophones, we would expect that we could locate the **rules** that determine which allophone occurs where. But a moment's reflection will show that there could be no such rules.

This is shown by the existence of **near-minimal pairs**, which can be defined as pairs which would be minimal except for some evidently irrelevant difference. Here are some near-minimal pairs for /ð / vs. /ʒ/:

tether	[ˈtɛðɚ]	vs.	*pleasure*	[ˈplɛʒɚ]
		or	*measure*	[ˈmɛʒɚ]
neither	[ˈniðɚ]	vs.	*seizure*	[ˈsiʒɚ]
lather	[ˈlæðɚ]	vs.	*azure*	[ˈæʒɚ]
heathen	[ˈhiðən]	vs.	*adhesion*	[ədˈhiʒən]
smoothen	[ˈsmuðən]	vs.	*illusion*	[ɪˈluʒən]
		or	*intrusion*	[ɪnˈtɹuʒən]
		or	*fusion*	[ˈfjuʒən]

This list shows that the phonetic environment has nothing to do with whether [ð] or [ʒ] occurs—there is no *consistent* factor that could determine which phone appears. Any effort to find the rules that determine the appearance of [ð] vs. [ʒ] would have to make use of a completely arbitrary collection of "environments" for these phones. If the rules cannot be found, then an analysis that claims that [ð] and [ʒ] are allophones cannot be justified.

It is also easy to imagine that if a new word came into English that created a true minimal pair (say, '*hesion*' to go with *heathen*), such a word would readily be accepted. It is logical, then, to assume that /ð/ and /ʒ/ are separate phonemes, and that (for some speakers) no fully minimal pairs happen to be available. The near-minimal pairs suffice to show this.

Plainly, the near-minimal pair method of establishing phonemes requires more work than minimal pairs do: it is the accumulation of forms, and the ruling out of all reasonable hypotheses concerning allophone environments, that ultimately permits near-minimal pairs to be used as evidence.

10.3 Using local environments to establish complementary distribution

The methods of minimal and near minimal pairs are used to establish that two sounds belong to separate phonemes. For establishing that two sounds are in the same phoneme, we need to

establish that they are in complementary distribution, and therefore we need to find the environments in which they occur. For this purpose, it is often useful to follow the method of **compiling local environments**, illustrated below.

The language we will examine is Maasai (Nilotic, spoken in Kenya and Tanzania), and our focus is solely on the following set of sounds: [p,t,k,b,d,g,β,ð,ɣ]. The last three of these are voiced fricatives: [β] is bilabial, [ð] dental, and [ɣ] velar. Below are 63 words containing these sounds.

Maasai data

1. [ailap] ‘to hate’
2. [aret] ‘to help’
3. [arup] ‘to heap up’
4. [asip] ‘to speak truly’
5. [ɓarːiɣoi] ‘reddish brown’
6. [ɓaða] ‘dangerous’
7. [ɗalut] ‘mischievous’
8. [ɗiɣai] ‘elsewhere’
9. [ɗorːop] ‘short’
10. [emɓiðir] ‘female wart hog’
11. [emaŋaða] ‘warriors’ village’
12. [embiɟan] ‘bravery’
13. [emburuo] ‘smoke’
14. [enɗaraða] ‘thunder’
15. [enɗuβai] ‘sisal’
16. [eŋɠirut] ‘silent-feminine’
17. [eŋɠoː] ‘small chest’
18. [enaiβoʃa] ‘Naivasha Lake’
19. [endaːraða] ‘fight each other’
20. [endorop] ‘bribe him’
21. [endulelei] ‘sodom apple’
22. [enduβeiðai] ‘Taveta woman’
23. [eŋgamaniɣi] ‘name of age-set’
24. [eŋgila] ‘garment-diminutive’
25. [eŋgiruðoðo] ‘fright’
26. [eŋgoː] ‘advise him’
27. [eŋoɣi] ‘sin’
28. [ilarak] ‘murderers’
29. [ilkeːk] ‘trees’
30. [ilpaβit] ‘hairs’
31. [iltoːi] ‘barrel’
32. [imɓok] ‘you detain’
33. [imbala] ‘papers’
34. [imbaɣiβak] ‘you are restless’
35. [imbok] ‘you clean ceremonially’
36. [indai] ‘you-plural’
37. [ijːoːk] ‘we’
38. [kaɣe] ‘but’
39. [kedʲanje] ‘left side’
40. [keβer] ‘heaven’
41. [kiɓiroðo] ‘stunted’
42. [koɣoː] ‘grandmother’
43. [oldʲiret] ‘pack saddle’
44. [olɗuɣa] ‘shop’
45. [olɠilaða] ‘room’
46. [olɟiβet] ‘stake’
47. [olkila] ‘garment’
48. [olkiɣuei] ‘thorn’
49. [olporːor] ‘age set’
50. [olpul] ‘slaughtering place’
51. [olpurɗa] ‘meat preserved in fat’
52. [olpurkel] ‘dry steppes’
53. [oltaː] ‘lamp’
54. [oltulet] ‘gourd in natural state’
55. [oltuli] ‘buttock’
56. [paɗan] ‘skilled in shooting’
57. [poɣira] ‘all’
58. [pus] ‘light colored’
59. [sarkin] ‘intermarriage taboo’
60. [taruɓini] ‘binoculars’
61. [tasat] ‘disabled’
62. [tisila] ‘sift it’
63. [tiʃila] ‘scrutinize it’

The first thing to notice about the data is that they include an additional series of stops: /ɓ ɗ ƒ ɠ/. These are voiced implosives, made by lowering the larynx to form a slight vacuum in the mouth during closure. The implosives form a separate series of phonemes in Maasai, as can be shown by minimal and near-minimal pairs such as the following:

35.	[im**b**ok]	'you clean ceremonially'
vs. 32.	[im**ɓ**ok]	'you detain'
7.	[**ɗ**alut]	'mischievous'
vs. 61.	[**t**asat]	'disabled'
26.	[eŋ**g**oː]	'advise him'
vs. 17.	[eŋ**ƒ**oː]	'small chest'
41.	[ki**ɓ**iroðo]	'stunted'
vs. 40.	[ke**β**er]	'heaven'

Having established this, we will ignore the implosives henceforth.

The method of compiling local environments works as follows: for each sound, we construct a list of all its appearances, each time including the preceding segment, if any, and the following segment, if any. For example, word #5, [ɓarːiɣoi] 'reddish brown' contains the target sound [ɣ]. This sound is preceded by [i] and followed by [o]. Thus, we add to our chart the following entry, in a column headed [ɣ]:

[ɣ]

/ i ___ o (5)

This chart entry may be read '[ɣ] occurs where preceded by [i] and followed by [o], in example (5).'

Where the target sound is the initial or final segment in the word, one includes in the environment a bracket of the type $]_{word}$ to designate this environment:

[k]

/ oː ___ $]_{word}$ (37)

This expression may be read "[k] occurs where preceded by [oː] and word-final."

One then continues through the whole set of data in this way. If this is done for the **velar** sounds only, one gets the following:

[k]		[g]		[ɣ]	
/ [$_{word}$ ___ a	(38)	/ ŋ ___ a	(23)	/ a ___ e	(38)
/ [$_{word}$ ___ e	(39, 40)	/ ŋ ___ i	(24, 25)	/ a ___ i	(34)
/ [$_{word}$ ___ i	(41)	/ ŋ ___ o	(26)	/ i ___ a	(8)
/ [$_{word}$ ___ o	(42)			/ i ___ i	(23)
/ l ___ e	(29)			/ i ___ o	(5)
/ l ___ i	(47)			/ i ___ u	(48)
/ l ___ i	(48)			/ o ___ i	(27, 57)
/ r ___ e	(52)			/ o ___ i	(57)
/ r ___ i	(59)			/ o ___ o	(42)
/ a ___]$_{word}$	(28)			/ u ___ a	(44)
/ a ___]$_{word}$	(34)				
/ e ___]$_{word}$	(29)				
/ o ___]$_{word}$	(32)				
/ o ___]$_{word}$	(35)				
/ oː ___]$_{word}$	(37)				

At this point, one inspects the data in hopes of locating general patterns. For these data, notice that [g] may occur only when the sound [ŋ] immediately precedes it. Further, and crucially, the sounds [k] and [ɣ] are **never** preceded by [ŋ]—which makes the distribution complementary. It thus looks likely that [g] is just one allophone of a phoneme, because it has such a highly restricted distribution. The preceding [ŋ] is likely to be the context that requires this allophone.

Inspecting the third column, we see another particular property: all cases of [ɣ] are surrounded by vowels. As before, this is not the case with the other candidate phones. The pattern suggests that [ɣ] is another allophone of the phoneme that includes [g].

Inspection of the [k] column shows no particularly interesting property: [k] may occur initially, after [r] or [l], and in final position. The only really important property here is that these various environments do **not** include the environments for [ɣ] or [g]. This makes [k] a good candidate for being an "elsewhere" allophone, in the sense described on p. 28 above.

We have established, then, that [k], [ɣ], and [g] are in complementary distribution: none occurs when any of the others may occur. The environments are shown below.

[g] / ŋ ___
[ɣ] / V ___ V where V stands for any vowel
[k] / elsewhere

It is reasonable to suppose that [k], as the elsewhere allophone, is the normal, unperturbed member of the phoneme, which we set up as the underlying representation. [g] and [ɣ] are

particular allophones resulting from phonological rules applying in particular environments. Note finally that, once we write the rules, the changes that they will carry out are not drastic: [k] and [g] differ only in voicing, while [k] and [ɣ] differ only in voicing and manner. This gives some additional plausibility to the idea that these sounds are related by rule.

Pursuing this, we can state the phonological analysis of these sounds as follows. First, /k/ is assumed to be a phoneme of Maasai, which undergoes the following two phonological rules.

/k/ Spirantization

k → ɣ / [+vowel] ___ [+vowel]

/k/ *is realized as* [ɣ] *between vowels*.

Postnasal Voicing

k → g / ŋ ___

/k/ *becomes voiced after* [ŋ].

Sample phonological derivations for three representative words of Maasai are as follows:

‘grandmother’	‘garment-dim.’	‘trees’	
/koko:/	/eŋkila/	/ilke:k/	underlying forms
ɣ	—	—	/k/ Spirantization
—	g	—	Postnasal Voicing
[koɣo:]	[eŋgila]	[ilke:k]	surface forms

It should be clear why the method of collecting local environments was useful here: as it turned out, the environments for the rules were in fact local, involving adjacent segments. While this is not true of all phonological rules,[23] it is common enough to make the strategy worthwhile.

Plainly, the procedure is tedious. It is possible, for many people and in many cases, to skip steps. If you have a knack for this, phoneme problems can be solved by inspection, without the tedious charting of all environments.

Another asset in solving such problems is experience. The same rules often show up in many different languages, so someone who has examined extensive phonological data before has a leg up in solving new problems. In the present case, we can note some languages that realize /k/ as [ɣ] between vowels: Taiwanese, Ewe (Ghana), and Tümpisa Shoshone (Death Valley, California). Languages that realize [k] as [g] after a nasal include Modern Greek, Leurbost

[23] The primary exceptions are vowel-to-vowel rules (example on p. 164), stress rules (treated in ch. 14), and the occasional long-distance consonant-to-consonant rule (see p. 89).

Gaelic (Scotland), and Waorani (Amazon basin, Peru). The questions of *why* the same rules recur in many different languages is one of the outstanding issues in phonological theory.

10.4 More Maasai: natural classes

We have not yet considered six of the nine Maasai sounds we set out to analyze, namely [p, b, β] and [t, ð, d]. Before proceeding, it is useful to arrange the relevant sounds into phonetic charts. Ideally, we would do this for all of the sounds of Maasai, but for present purposes the following will suffice:

	Bilabial	*Dental*	*Velar*
voiceless stops	p	t	k
plain voiced stops	b	d	g
voiced implosive stops	ɓ	ɗ	ɠ
voiced fricatives	β	ð	ɣ
voiced nasals	m	n	ŋ

If we sort out the target sounds in the way we did before, we will get the following:

[p]

/ [$_{word}$ ___ a (56)
/ [$_{word}$ ___ o (57)
/ [$_{word}$ ___ u (58)
/ l ___ a (30)
/ l ___ o (49)
/ l ___ u (50, 51, 52)
/ a ___]$_{word}$ (1)
/ i ___]$_{word}$ (4)
/ o ___]$_{word}$ (9, 20)
/ u ___]$_{word}$ (3)

[b]

/ m ___ a (33, 34)
/ m ___ i (12)
/ m ___ o (35)
/ m ___ u (13)

[β]

/ a ___ i (30)
/ e ___ e (40)
/ i ___ a (34)
/ i ___ e (46)
/ i ___ o (18)
/ u ___ a (15)
/ u ___ e (22)

[t]

/ [$_{word}$ ___ a (60, 61)
/ [$_{word}$ ___ i (62, 63)
/ l ___ a (53)
/ l ___ o (31)
/ l ___ u (54, 55)
/ a ___]$_{word}$ (61)
/ e ___]$_{word}$ (2, 43, 46, 54)
/ i ___]$_{word}$ (30)
/ u ___]$_{word}$ (7, 16)
/ u ___]$_{word}$ (16)
/ u ___]$_{word}$ (16)

[d]

/ n ___ a (19, 36)
/ n ___ o (20)
/ n ___ u (21, 22)

[ð]

/ a ___ a (6, 11, 14, 19, 45)
/ i ___ a (22)
/ i ___ i (10)
/ o ___ o (25, 41)
/ u ___ o (25)

If you consider both the phonetic chart and the list of environments, you can see that the distribution of the bilabial and dental sounds is in complete parallel with the velars: voiced stops appear after nasal consonants, voiced fricatives occur between vowels, and voiceless stops occur elsewhere.

Thus, although we are dealing with three phonemes and nine allophones, we do not need a large number of rules to cover the data. Rather, we can use features to write general rules that cover all three phonemes at once. The specific analysis sets up the three phonemes /p/, /t/, and /k/ and posits two generalized phonological rules.

Spirantization

$$\begin{bmatrix}+\text{stop}\\ -\text{voice}\end{bmatrix} \rightarrow \begin{bmatrix}+\text{voice}\\ -\text{stop}\\ +\text{fricative}\end{bmatrix} / \text{[+vowel]} \_\_\_ \text{[+vowel]}$$

A voiceless stop is realized as the corresponding voiced fricative when surrounded by vowels.

Postnasal Voicing

[+stop] → [+voice] / [+nasal] ___

A voiceless stop is realized as the corresponding voiced stop when it follows a nasal consonant.

For this approach to work, we need to be explicit about how features are used in rules. If a feature occurs on the right side of the arrow, that feature is changed, whenever the rule applies. But *all other features are assumed to remain unaltered*. Thus, if we are considering a sequence like /mp/ and apply Postnasal Voicing (as in #33, /impala/ → [im**b**ala]), the [–voice] of the /p/ is changed to [+voice], so that /p/ is altered to [b]. But the features [+bilabial] and [+stop] remain unaltered. In this way, we can express rules that alter whole classes of segments (such as all the voiceless stops) in parallel. The features therefore permit a simpler and more general analysis than would be available if all the allophones of each phoneme were derived separately.

The fact that the stop phonemes of Maasai vary in parallel fashion is not an accident. The same phenomenon shows up in a great number of languages. Here are two examples we've already covered of how rules apply to classes of sounds.

Vowel Shortening in English: The shortening of /eɪ/ to [ĕɪ] before voiceless consonants in English (p. 30) is not unique to /eɪ/: all vowels of English are shortened in this environment. Examples: *coat* [kŏʊt] vs. *code* [koʊd], *lap* [lăp] vs. *lab* [læb], etc.

Spirantization in Spanish: Spanish not only has [ð] as a post-vowel allophone of /d/ (p. 37), but also [β] as a post-vowel allophone of /b/ and [ɣ] as a post-vowel allophone of /g/. In

other words, all voiced stops are converted to the corresponding fricatives in the post-vowel environment. Examples: /ˈlobo/ [ˈloβo] 'wolf', /ˈlago/ [ˈlaɣo]'lake'.

The general lesson that we learn from these examples (and countless others) is this: *phonological rules are based on phonetic features*. This general principle has three specific subcases.

First, the *set of sounds a rule applies to* is normally a set of sounds that share a particular phonetic feature or set of features. For example, the Spirantization rule of Spanish applies to all and only the voiced stops, characterized as [+stop, +voice].

Second, rules often change only *one or two features* of a sound, rather than making massive alterations. For example, the rules for Maasai alter only voicing and the stop/fricative distinction.

Lastly, the *sounds appearing in the environment of a rule* are almost always a set of sounds that share a particular phonetic feature or features. For example, the rule of English that shortens vowels applies before the complete set of consonants in English that are [–voice].

A **natural class** of sounds is defined as any complete set of sounds in a given language that share the same value for a feature or set of features. For example, /m/, /n/, and /ŋ/ in Maasai and in English form a natural class because they constitute the complete set of sounds that share the feature [nasal]. Likewise, /p/, /t/, and /k/ form a natural class in Maasai and in English because they constitute all the [+stop,–voiced] sounds of the language.

It can be noted that that the natural class defined by a particular feature combination will vary from language to language, simply because different languages have different inventories of sounds. Thus, in English [p t k] form the natural class of voiceless stops ([+stop, –voice]). Yet [p t k] are not a natural class in Persian (Farsi), since Persian contains a fourth voiceless stop, uvular [q]. For Persian, [+stop, –voice] is a natural class, but consists instead of the sounds [p t k q].

To reiterate the point made above with the novel terminology: in most instances, the segments that undergo a rule or appear in the environment of a rule form a natural class in the language in question.

11. Exercises

11.1 /ɹ/ in American English

This is a simple allophone problem, to be solved like the Maasai allophone problem in §10.3 above. The focus sounds are the voiced alveolar central approximant [ɹ], and the (slightly) rounded voiced alveolar central approximant [ɹʷ].

migrants	[ˈmaɪgɹʷənts]	*Homeric*	[hoʊˈmɛɹʷɪk]
or	[ˈɔɹ]	*trek*	[ˈtɹʷɛk]
from	[ˈfɹʷʌm]	*debriefed*	[diˈbɹʷift]
shire	[ˈʃaɪɹ]	*reply*	[ɹʷiˈplaɪ]
tripling	[ˈtɹʷɪplɪŋ]	*Iraqi*	[ɪˈɹʷɑki]
metaphor	[ˈmɛtəˌfɔɹ]	*preys*	[ˈpɹʷeɪz]
iridium	[ɪˈɹʷɪdiəm]	*ranted*	[ˈɹʷæntəd]
proclivities	[pɹʷoʊˈklɪvəɾiz]	*crucible*	[ˈkɹʷusəbəl]
romancing	[ɹʷoʊˈmænsɪŋ]	*indiscriminately*	[ˌɪndəsˈkɹʷɪmənətli]
February	[ˈfɛbjuɛɹʷi]	*fear*	[ˈfɪɹ]
dwarfing	[ˈdwɔɹfɪŋ]	*dreadful*	[ˈdɹʷɛdfəl]
assure	[əˈʃuɹ]	*feldspar*	[ˈfɛldspɑɹ]

11.2 phonemes

Lango is a Nilotic language spoken in Uganda.

a. Make a phonetic chart of all the consonants in the data below (columns: place of articulation; rows: manner of articulation, voicing, and length).

b. This problem deals just with the sounds [p, pp, ɸ, t, tt, ɾ̥, t͡ɕ, tt͡ɕ, ɕ, k, kk, x]. Collect local environments for these sounds only, following the method given in this chapter.

c. The sounds [p, pp, ɸ, t, tt, ɾ̥, t͡ɕ, tt͡ɕ, ɕ, k, kk, x] may be grouped into eight phonemes. List the eight phonemes and their allophones. State the environments where the allophones occur. You may use "elsewhere" to simplify your presentation.

d. State in words the phonological rules that determine the allophones. There is a major ambiguity in determining what is an allophone of what. Figure out this ambiguity, and state analyses for both possibilities.

e. The word for 'lazy' is [ɲàp], with a [p]. The word for 'laziness' is [ɲáɸô], with a [ɸ]. Explain how this bears on the ambiguity noted in the previous question.

Phonetic symbols:

- [á] marks High tone, [à] marks Low, [â] marks Falling.
- [t͡ɕ] and [d͡ʑ] are alveolopalatal affricates, [ɕ] is an alveolopalatal fricative.
- [ɾ̥] is a voiceless tap.
- Consonants transcribed as double are simply held longer; they are not "rearticulated". Think of them as single long consonants.

1	[pì]	'because of'	33	[dáxô]	'woman'
2	[kèt͡ɕ]	'hunger'	34	[t͡ɕùt͡ɕ]	'pitch black'
3	[tóŋ]	'spear'	35	[tódd͡ʑó]	'to beat up'
4	[búttt͡ɕó]	'to yell at'	36	[wókkí]	'a few minutes ago'
5	[t͡ɕɔ̀ː]	'men'	37	[dìəxə̀]	'wet'
6	[ʔɔ̀t]	'house'	38	[máxât͡ɕ]	'scissors'
7	[dɔ̀ttɔ̀]	'to suck'	39	[pé]	'snow, hail'
8	[pə̀ppì]	'fathers'	40	[kɔ́ppɔ̀]	'cup'
9	[pójó]	'to remember'	41	[pàtt͡ɕó]	'to peel'
10	[ljèt]	'hot'	42	[pámmà]	'cotton'
11	[bókkó]	'to make red'	43	[mɔ̀ɾ̥ɔ̀xà]	'car'
12	[jútt͡ɕú]	'to throw'	44	[bə̀p]	'to deflate'
13	[èŋə́ɾ̥ó]	'lion'	45	[lwìttê]	'to sneak'
14	[ókkɔ́]	'completely'	46	[ɲàp]	'lazy'
15	[déɸô]	'to collect'	47	[bwɔ̀ttɔ̀]	'to retort insultingly'
16	[dɛ̀k]	'stew'	48	[tèttó]	'to forge'
17	[t͡ɕùɸâ]	'bottle'	49	[t͡ɕàmmó]	'to eat'
18	[gwɛ̀k]	'gazelle'	50	[tɔ̀p]	'to spoil'
19	[kókkó]	'to cry'	51	[t͡ɕɔ́k]	'near'
20	[ɲáɸô]	'laziness'	52	[pàɸó]	'father'
21	[ɾét͡ɕ]	'fish'	53	[ŋwètt͡ɕó]	'to run from'
22	[bóɾ̥ə́]	'to me'	54	[bót]	'to'
23	[dìppó]	'to smash'	55	[dɛ̀ppó]	'to collect'
24	[dwéɾ̥ê]	'months'	56	[gɔ̀t]	'mountain'
25	[kóddó]	'to blow'	57	[jìt͡ɕ]	'belly'
26	[t͡ɕín]	'intestines'	58	[bìttó]	'to unshell'
27	[gíɾ̥é]	'really'	59	[dɔ̀k]	'to go back'
28	[lòɕə̀]	'man'	60	[kòp]	'matter'

29	[kwə̀ɕê]	‘leopards’	61	[tîn]	‘today’
30	[kál]	‘millet’	62	[kít]	‘kind’
31	[màɕê]	‘fires’	63	[àkká]	‘purposely’
32	[àbíɕèl]	‘six’	64	[t͡ɕàk]	‘milk’

Further Reading

The opening of this chapter states that the central subject matter of phonology is *sound patterns* in language. This invokes two important early phonological works. Edward Sapir's "Sound patterns in language" (1925; *Language* 1:37-51) was the first work to point out that two languages could have phonetically identical inventories but quite different phonologies (see §9 above). *The Sound Pattern of English*, by Noam Chomsky and Morris Halle (1968, Harper and Row) is by consensus the most important single work in phonological theory. Many of the ideas given in this text first appeared there.

The systematization of a procedure for finding the phonemes of a language was one of the major accomplishments of the so-called "American structuralist" school of linguistics, which flourished from approximately the 1920's to the 1950's. A fine presentation of the method of phonemicization by a member of this school may be found in H. A. Gleason's *An Introduction to Descriptive Linguistics* (1961, Holt, Rinehart and Winston). Two works that are widely considered to be gems of American structuralism are *Language* by Leonard Bloomfield (1933; reprinted 1984; University of Chicago Press) and *Language: An Introduction to the Study of Speech* by Edward Sapir (1921, Harcourt Brace; now on line at http://www.bartleby.com/186/).

The maximum and minimum phoneme counts in §2 are taken from Ian Maddieson's *Patterns of Sounds* (1984, Cambridge University Press), a very useful survey of several hundred phoneme inventories.

The point that shortening of vowels before voiceless consonants is a rule specific to particular languages is argued for in Patricia Keating "Universal phonetics and the organization of grammars," in Victoria Fromkin, ed., *Phonetic Linguistics* (1985, Academic Press).

Maasai phonemes: Archibald N. Tucker and J. Tompo Ole Mpaayei, *A Maasai Grammar with Vocabulary* (1955, Longman, Green).