1 The Sounds of Speech

This chapter has several goals. First, it surveys the sounds of speech from a traditional phonetic point of view, introducing the customary categorizations for the most common sounds in the world's languages. Second, it illustrates the basic notion of phonological natural classes and motivates the distinctive feature notation to express such classes. It then systematically surveys the universal phonetic alphabet, introducing a set of features to serve as the notational foundation for the analytic concepts and techniques of chapters 2 and 3. Finally, it familiarizes the reader with the most common systems of phonetic transcription.

In its desire to give a reasonably complete first-pass survey of the phonetic alphabet, this chapter is necessarily lengthy. Readers wishing to get to the phonological action sooner may content themselves with reading the first five or six sections. Those desiring more background should consult a phonetics text such as Ladefoged 1993.

1.1 The Saussurean Sign

According to Miller and Gildea (1987), in the normal course of development a child learns a vocabulary of some 80,000 lexical items. Many adults have lexicons of much greater size. Lexical acquisition is the one most obvious aspect of language development that requires a large contribution of memorization and learning. At the level of discussion relevant here, we conceive of the lexical item as a Saussurean sign. Our discussion focuses on the *signifiant* (*phonological form* in generative parlance). To build a large vocabulary, the learning system must have the capacity to readily perceive the information constituting the *signifiant* in the linguistic environment, to store and recall it at will, and to articulate it through a channel that offers easy transmission. The natural mode of articulation for human language is speech (though when the auditory system is impaired, language is readily expressible in the visual modality of sign language). To understand some of its most basic properties, we can ask what design features it would be desirable to build into a language articulation system so that it may subserve vocabulary acquisition. Surely one requirement is that the system be able to assign different lexical items distinct articulations that are detectable by the human ear. Second, the faster the item can be articulated and the more quickly the information contained in the acoustic signal can be recovered, the better. Third, the system itself must be easy to learn and to operate.

While a groan and a giggle are readily distinguished from one another, no language encodes its vocabulary in such gross vocal wholes. Think of the cognitive hardware that would be required to perceive, learn, and put into operation 80,000 different gestures. A much more efficient system would stipulate a small number of basic atoms and some simple method for combining them to produce structured wholes. For example, two iterations of a concatenation operation on an inventory of 10 elements (say, the digits 0,1,2,3,4,5,6,7,8,9) will distinguish 10^3 items (e.g., 000, 001, 002, . . . , 999). As a first approximation, it can be said that every language organizes its lexicon in this basic fashion. A certain set of speech sounds is stipulated as raw material. Distinct lexical items are constructed by chaining these elements together like beads on a string. Two lexical items are *distinct* if they differ in length (e.g., *us, bus, bust, burst, bursts*) or have different elements at any given position in the string (e.g., *bus, sub, rub, urb*). The fact that all languages studied so far submit to an alphabetic representation and that such writing systems are learned with relative ease reflects this basic mode of linguistic organization. However, languages typically utilize more than ten basic elements. Also, while the digits {0,1,2, . . . , 9} can be combined in any order to produce a code, human languages impose severe constraints on possible sound sequencing. Phonology can be defined as the attempt to reach a deeper understanding of the nature of these constraints.

Some constraints clearly reflect exigencies of the articulatory hardware. For example, longish sequences of consonants are not preferred. Too many unreleased consonants shut off airflow and transmission ceases. As frequently observed, the human vocal apparatus did not evolve specifically in order to articulate language. The lungs, tongue, and lips were present in the species well before the advent of language, which has had to make do with these anatomical givens (though see Lieberman 1984 for arguments that certain evolutionary developments such as the lowering of the larynx have accommodated the language faculty). But while some features of phonological structure are explicable in mechanical terms, many others are not. There appear to be principles, laws, tendencies, and modes of organization that are more cognitively based. Their discovery, study, and formulation is what animates much of the interest in the subject.

1.2 Speech Sounds: An Internal Dimension

Lexical items (*dog, cat, man*) resolve themselves into strings of speech sounds ([dɔg], [kæt], [mæn]). In addition to this "horizontal" dimension, the individual speech sounds themselves submit to further analysis. They are not the atoms of linguistic organization but constellations of still more fundamental categories. Study of the sound systems of the many different languages described over the past two centuries strongly suggests that the sounds of speech are drawn from a tightly constrained, *universal phonetic alphabet* with a surprisingly rich internal structure. This structure in large part reflects categorizations imposed by the vocal apparatus.

1.2.1 The Vocal Apparatus and the Articulation of Consonants

We can get a sense of this inner dimension of linguistic organization by a brief review of the gross structure of the vocal apparatus (see chapter 4 and figures 4.1 and 4.5 for more extensive discussion). Human speech has often been described as "movements made audible by the vocal apparatus." Between the time that air is first expelled from the lungs and the time it finally exits the oral and nasal cavities, it is excited and modified by the organs of speech in specific ways to which the auditory system is sensitive. Recent research suggests that after air has been expelled from the lungs into the trachea or windpipe, six separate *articulators* may modify it in linguistically significant ways. These articulators are the *larynx,* the *tongue root,* the *velum,* the *tongue body,* the *tongue blade,* and the *lips.*

The larynx houses the vocal cords (folds) – two elastic lips that form a valve sitting across the glottis. The vocal folds can assume a number of configurations that are linguistically significant. Here we distinguish two. First, the folds may be brought together (*adducted*) along their entire length in such a way as to be set in vibration when air passes between them. Sounds produced with this laryngeal configuration are called *voiced.* They are opposed to *voiceless* sounds, which lack vibration due to an instruction to either separate the folds or increase their tension. For example, the initial [s] and [z] sounds of *sip* and *zip* are opposed as voiceless to voiced.

After the airstream leaves the larynx, it passes through as many as three cavities whose specific properties excite it in particular ways. These are the *pharyngeal, nasal,* and *oral cavities.* In the pharynx, the root of the tongue may be projected forward to create a greater pharyngeal opening or it may fail to be so advanced. This advancement of the tongue root is the basis of the distinction between the relatively tense-feeling, crisper vowels of *beat* and *bait* and the more lax vowels of *bit* and *bet.* Prolongation of the vowel in *beat* produces the sensation of a noticeable tensing of the tongue muscles; the tension reflects the advancement gesture. As we will see, *advancement* of the tongue root (ATR) is significant not only in English but also in many African languages, which have a principle of vowel harmony that prohibits the mixture of advanced and nonadvanced vowels within a word.

After the airstream reaches the upper portion of the pharyngeal cavity, it may be expelled entirely into the oral cavity or part of it may enter the nasal cavity by lowering of the velum (the back portion of the roof of the mouth, sometimes also called the *soft palate*), which acts as a valve shutting off the nasal passage. Sounds produced with a lowered velum are called *nasal;* those produced with the velum in a raised position are called *oral.* The distinction between the raised and lowered velum position is what distinguishes the English words *bat, dip* from *mat, nip.* Virtually all languages distinguish at least one nasal consonant from a corresponding oral one. Vowel sounds may also be either oral or nasal. In the speech of most English speakers, the vowel of *can't* is noticeably nasalized as compared to the vowel of *cat.* In some American dialects the [n] is deleted, so that the nasalization in the vowel is the only feature distinguishing *can't* from *cat.*

The three remaining articulators are housed in the oral cavity. From front to back, they are the lips, the front portion (blade) of the tongue, and the tongue body. The term *labial* designates sounds articulated by the lips. Sounds articulated with the muscles controlling the front portion of the tongue are known as *coronals*, and those articulated by the tongue body are termed *dorsal* or *velar*. Labial sounds include the initial consonants in *pit, bit,* and *mitt;* the final consonants in *fat, fad,* and *fan* are coronals; and *tack, tag,* and *tang* end in velar consonants.

So far, each of the phonetic dimensions we have looked at is localized in a particular articulator. For example, the voicing dimension is a function of the larynx, and nasality is a function of the velum. Degree of stricture is a phonetic category that cuts across several different articulators. *Occlusive* or *stop* consonants are formed by having the articulator produce a closure that temporarily interrupts the flow of air. They are opposed to *fricatives,* in which the articulator forms a stricture that is narrow enough to create air turbulence but not so narrow as to interrupt the airflow. For example, the stop [p] of *pile* is produced by making a closure at the lips; in the fricative [f] of *file,* the lower lip touches the upper teeth but the contact is not sufficient to block airflow past the teeth. Similarly, in the stop [t] of *tin,* the front of the tongue makes a closure sufficient to interrupt airflow, while in the [s] of *sin,* the tongue is positioned with respect to the teeth and gum ridge so as to create turbulence but not to interrupt airflow. The initial labials in *bile* and *vile* and the coronals in *dip* and *zip* exemplify the same stop versus fricative opposition – but combined with a voicing of the vocal folds. While not bound to a particular articulator and thus somewhat more abstract, the stop-fricative distinction is nevertheless an important one in many languages. For example, Hebrew has a rule that relates stops and fricatives in particular phonological contexts: compare the perfect and imperfect forms of the verbs [p]*agaš, yi-*[f]*goš* 'meet'; [b]*aḥar, yi-*[v]*ḥar* 'choose'; and [k]*atab, yi-*[x]*tob* 'write'. It is interesting that such stop-fricative pairs as *p-f, b-v, k-x* are spelled by the same letter in the Hebrew alphabet. Whether the letter stands for a stop or a fricative is indicated by context as well as by the presence versus absence of a diacritic (a dot known as the *daghesh*). Recourse to this diacritic reveals an intuitive facility for resolving speech sounds into their constituent features.

The consonants made available by the various articulatory categories surveyed above are tabulated in (1). They are designated by the customary phonetic symbols. The only ones not found in English are the velar fricatives [x] and [ɣ]. The former occurs as the final consonant in German *Bach* and the latter is the phonetic realization of the second consonant in Spanish *lago* 'lake'.

(1)

	labial	coronal	velar
voiceless stops	p	t	k
voiced stops	b	d	g
voiceless fricatives	f	s	x
voiced fricatives	v	z	ɣ
nasals	m	n	ŋ

Several points are worth noting. The first is a combinatoric one. In terms of its articulation, every speech sound in (1) is the product of a specific choice for each of the phonetic gestures introduced above: Are the vocal folds brought together or separated? Is the velum raised or not? Does the stricture interrupt the airflow? Thus, [p] shares the stop property with [b] and [k]. It shares the labial property with [f] and [v] and the voiceless property with [x] and [s]. From this perspective, we see that the phonetic gestures (the distinctive features) are the ultimate units of speech. Any particular sound is a macrolevel constellation of such properties – a bundle of distinctive features.

One might wonder, however, whether this phonetic reductionism has any linguistic relevance. After all, tables and chairs are composed of atoms and molecules; but we experience these objects as tables and chairs – not as atoms. Note first of all that cross-classification of the set of speech sounds into such a densely packed network is an efficient way to catalogue a large inventory of items (such as the vocabulary of a language). Just as the appropriate shoes in a shoestore can be quickly located because they are shelved in terms of a relatively small number of categories (adult vs. child, sex, length, width, color, style, etc.), so a large vocabulary can be deployed in speech if individual lexical items can be located quickly. From the table in (1), we have seen how just four parameters (place of articulation, voicing, nasality, and continuancy) yield fifteen distinct sounds. Furthermore, it is obviously significant that while the various phonetic gestures composing a sound can be produced simultaneously, the auditory system is still able to resolve the resultant complex sound into its constituent parts. The payoff is the increased efficiency with which messages can be transmitted (coded and decoded). Thus, while with just four phonetic parameters 15 distinct messages can be transmitted in one unit of time, 15 * 15 distinct messages are possible in two units of time. This mode of articulation makes it possible to transmit a greater number of messages per unit of time. Finally, since the phonetic dimensions comprising the universal phonetic alphabet are the same for all languages, the basic system is already "wired in" and need not be learned. What must be learned are which particular dimensions and combinations are employed in the language of the environment.

A few important qualifications to the combinatorics are necessary here. First, certain features are more compatible than others. Note from (1) that while voicing and continuancy (the stop vs. fricative dimension) combine freely at all three points of articulation, nasality is more restricted. There is a marked preference for voiced nasals. Although voiceless nasals can be found (e.g., in Burmese), they are decidedly rare. Also, nasality combines with oral closure in [m], [n], and [ŋ]. Nasal fricatives do not make for a happy combination of features – perhaps because so much airflow is diverted to the nasal cavity that not enough remains to generate the turbulence required of a fricative. Which feature combinations are optimal and why is still a poorly understood problem. The table in (1) represents most of the frequently occurring consonants. Later we will see how the system can be augmented to describe the fuller range of consonants found in the world's languages.

In generative grammar the major source of evidence for the features has been their utility in providing the natural cuts of phonetic space needed to understand phonological patterning. For example, recall from our discussion of *writer-rider* in the Introduction that vowels in English are shorter before [p,t,k,s] than before [b,d,g,z]. We now see that the basis of this difference is not haphazard or arbitrary; it can be expressed as a function of the voicing of the following consonant: vowels are short before the voiceless consonants [p,t,k,s] and long before the voiced [b,d,g,z]. Similarly, the English dialect in which *limp, lint,* and *link* are realized as [lĩp], [lĩt], and [lĩk] has a rule that deletes nasal consonants before voiceless stops. Before developing a more precise notation to express such natural classes, we will briefly survey the articulation of vowels.

1.2.2 Vowels

Vowels are distinguished from consonants primarily by a less radical degree of constriction imposed by the lips and tongue on the flow of air through the mouth. Distinctions within the class of vowels are created by the specific shape of the lips and the precise positioning of the tongue body. It is traditional to describe these tongue positions by reference to a neutral point such as that corresponding roughly to the location the tongue body occupies in pronouncing the vowel in the English word *bed.* (Just before the onset of speech, the vocal apparatus shifts to a "get-ready" configuration in which the vocal folds are in the voiced position, the velum is raised, and the tongue is positioned in the mid front region where the vowel of *bed* is articulated.) The vowel of *bit* is articulated by raising the tongue body above the neutral position of *bet.* A similar raising distinguishes the ATR vowel of *beat* from that in *bait.* The vowel in *bat* is produced by lowering the tongue below the neutral position. Finally, comparison of the tongue positions for the vowels in pairs such as *bet* vs. *bought* or *bait* vs. *boat* reveals a retraction from the neutral position. Vowels produced with a raising, lowering, and retraction of the tongue body are called *high, low,* and *back* vowels, respectively. While the tongue cannot be simultaneously raised and lowered, the retraction gesture combines freely with the instructions to raise, lower, or retain the tongue in the neutral position. Various English words whose vowels differ in relative tongue positioning in just these ways are displayed in (2). (The [+ATR] vowels of *beat, bait, food,* and *boat* are also longer than their [−ATR] counterparts.) The corresponding symbols that customarily designate these vowel qualities are indicated in (3). The high lax [−ATR] vowels are represented here by small capitals [ɪ] and [ᴜ]; the International Phonetic Association (IPA) utilizes the special symbols [ɩ] and [ɷ] instead.

(2)

	front		back	
	[+ATR]	[−ATR]	[−ATR]	[+ATR]
high	beat	bit	foot	food
mid	bait	bet	bought	boat
low		bat	Bach	

(3)

	i	ɪ	ᴜ	u
	e	ɛ	ɔ	o
		æ	ɑ	

1.3 Natural Classes

After this brief phonetic review, we begin looking at some of the phonological properties of speech sounds. One of the basic insights of 20th-century linguistics is that an individual sound tends to pattern with certain other sounds in the overall fabric of any given language. Such sound groupings are termed *natural classes.* Our first illustration comes from Chamorro, an Austronesian language of Guam (Topping 1968). Chamorro distributes its vowels in front-back pairs over the three tongue heights (4). (In unstressed syllables the distinctions between [i] and [e], [u] and [o], [æ] and [a] are lost, a detail that we overlook here.) In Chamorro, the first vowel of certain words shows a systematic change when preceded by certain particles whose precise characterization is not important at this point (5).

(4)		front	back
	high	i	u
	mid	e	o
	low	æ	a

(5)				
	gumə	'house'	i gimə	'the house'
	tomʊ	'knee'	i temʊ	'the knee'
	lahɪ	'male'	i læhɪ	'the male'
	gwihən	'fish'	i gwihən	'the fish'
	pecʊ	'chest'	i pecʊ	'the chest'

In (5) we see that a noun such as *gumə* 'house' whose first syllable contains [u] replaces that vowel with [i] when preceded by the definite particle *i*. In a similar way, [o] alternates with [e] in *tomʊ* 'knee' and [a] with [æ] in *lahɪ* 'male'. The vowels [i,e,æ] fail to change in this environment; only [u,o,a] alternate. Given the six-element inventory of (4), twenty possible three-member subsets can be formed. But only a very small percentage ever figure in the grammar of a language in the systematic way in which [u,o,a] function in Chamorro. Other logically possible drawings such as [i,o,æ] or [u,e,a] seldom appear in any rule in any language. This is a remarkable and significant fact that any theory of phonology must explain.

To continue our discussion of natural classes, the definite particle *i* is not the only one that activates the Chamorro vowel alternation. The examples in (6) show that *en, sæn, mi,* and *gi* do as well.

(6)				
	tunʊʔ	'to know'	en tinʊʔ	'you know'
	hulʊʔ	'up'	sæn hilʊʔ	'upward'
	otdʊt	'ant'	mi etdʊt	'lots of ants'
	oksʊʔ	'hill'	gi eksʊʔ	'at the hill'
	lagʊ	'north'	sæn lægʊ	'toward north'

The vowels in these particles are restricted to [i,e,æ]. This set (the complement of [u,o,a]) is the only other three-member subset drawn from (4) that appears with any frequency in the world's languages. It too constitutes a natural class.

We have noted two ways in which natural classes figure in phonological alternations: in defining the set of segments participating in the alternation, and in defining a conditioning environment. The individual pairings that comprise an alternation constitute a third way in which sounds organize themselves into natural groupings. For example, given the six-vowel inventory [i,e,æ,u,o,a], there are many logically possible mappings of [u,o,a] onto some other subset that could in principle define an alternation. However, only a very small number are attested empirically. Pretheoretically, some notion of "phonological distance" constrains the possible pairings. Segments are not equidistant in "phonological space." For example, [o]≈[u] is a frequent alternation; but the chances of encountering [a]≈[s] are virtually nil. Each pairing displayed in the Chamorro alternation ([u]≈[i], [o]≈[e], and [a]≈[æ]) is widely attested.

Finally, there is characteristically an internal coherence to the various pairings that comprise a phonological alternation. While [o]≈[u] is as natural a pairing as [o]≈[e] is, it is very unlikely that [o] will alternate with [u] in the same circumstances in which [a] alternates with [æ] and [u] with [i]. In a crucial sense, the pairing of [o] with [e] is of the same kind as the [u]≈[i] and [a]≈[æ] pairings.

To sum up, only certain restricted combinations of sounds are likely to constitute the various components of a phonological alternation: the *focus* or input (e.g., [u,o,a] in Chamorro), the *image* or output (Chamorro [i,e,æ]), and the *conditioning context* (Chamorro [i,e,æ] in a particle). "Natural phonological class" is not a recent discovery. Designations such as "voiceless stops" and "front vowels" appear in many of the sound laws discovered in the 19th century. The important question is, Why is human language organized in this fashion? Although "back vowel" is defined with respect to the neutral tongue position, the latter, so far as is known, has no independent motivation outside of language. The tongue does not naturally gravitate toward [ɛ], except in the articulation of human language. Consequently, the natural phonological classes must arise from and be explained by the particular way in which UG organizes the information that determines how human language is articulated and perceived. The real question then is the nature of this organization.

1.4 Distinctive Features

The Russian-born linguist Roman Jakobson (1896–1982) proposed an answer to this question that is generally regarded as one of the most important discoveries in the history of the discipline. Jakobson's proposal (based on collaboration with his compatriot Nikolai Trubetzkoy (1890–1938)) is that phonological segments can be analyzed into complexes of *distinctive features* that cross-classify the entire inventory of possible speech sounds into a densely packed network. In general, each feature comprehends two possible values, represented as plus or minus. Also, each feature has its own characteristic articulatory and acoustic correlate.

For example, the vowels of (2) receive the distinctive feature analysis in (7). (When the distinction between back [ɑ] and central [a] is not relevant, [ɑ] will be transcribed as [a].)

(7)

	i	ɪ	e	ɛ	æ	a	o	ɔ	u	ʊ
high	+	+	−	−	−	−	−	−	+	+
low	−	−	−	−	+	+	−	−	−	−
back	−	−	−	−	−	+	+	+	+	+
ATR	+	−	+	−	−	−	+	−	+	−

We may interpret [+high] as the instruction the brain sends to the vocal apparatus to raise the tongue body above the neutral point. Segments with the [−high] designation do not raise the tongue body. Sounds that are [+back] are articulated by retracting the tongue body from the reference point. [−back] defines a sound whose articulation lacks this tongue retraction. As a first approximation, each feature implies an independent phonetic dimension. Specification as plus or minus picks out a particular end point on the relevant dimension. The complete stock of features thus constitutes a hypothesis about the phonologically significant phonetic dimensions along which possible speech sounds can vary. An adequate system should be able to resolve any sound from any language into its constituent features. Jakobson's proposal thus encourages us to think of phonological segments such as [i] and [u] as the *feature matrices* in (8). Each such matrix indicates the linguistically significant configurations the vocal apparatus assumes in order to articulate the corresponding sound.

(8) $[\text{i}] = \begin{bmatrix} +\text{high} \\ -\text{low} \\ -\text{back} \\ +\text{ATR} \end{bmatrix} \qquad [\text{u}] = \begin{bmatrix} +\text{high} \\ -\text{low} \\ +\text{back} \\ +\text{ATR} \end{bmatrix}$

The proposal to represent phonological segments as feature matrices has a number of positive consequences for the issue of natural classes. First, it now becomes possible to formalize the notion of phonological distance in terms of shared feature specifications. Given the feature system in (7), [a] is closer to [æ] than [o] is, because [a] and [æ] differ by just one feature while [o] and [æ] differ in three features. Second, we can begin to understand why certain collections of segments are more natural than others. For example, the set [u,o,a] can be extracted from the Chamorro inventory [u,i,o,e,a,æ] by the simple designation [+back]. Similarly, the vocalic segments [i,e,æ] that constitute the environment for the Chamorro alternation can be specified as all and only the vowels of the language with the feature [−back].

We also note that there is no way in which such unnatural sets as [u,e,a] or [i,o,æ] can be uniquely identified as a conjunction of features. In order to specify exactly the set [u,e,a] and leave out [i,o,æ], several disjunctive statements are required. For example, while [u,a] are back vowels, [+back] also includes [o] and excludes [e]. To let in [e] but leave out [o] necessitates a complicated statement such as "[−back] only if [−high, −low] and otherwise [+back, +high] or

[+back, +low]." But this is precisely the result we seek. If natural classes have simple analyses and unnatural ones do not, then the theory is drawing the empirically correct distinctions and offers a formal basis for explaining, as opposed to simply describing, why things are as we appear to find them.

Finally, we can reconstruct the notion of the internal coherence of an alternation by noting the extent to which the feature differences between the two terms [a]≈[b] in one alternating pair are matched by corresponding differences in the other alternating pairs [c]≈[d], and so on. Given the feature system in (7), we see that the Chamorro pairings [u]≈[i], [o]≈[e], and [a]≈[æ] reflect the same feature change of [+back] to [−back].

The payoff is that we can formalize the Chamorro sound change by the rule in (9). (We assume a feature [±consonantal] to distinguish [+consonantal] consonants from [−consonantal] vowels; see section 1.8.)

(9) $$\begin{bmatrix} -\text{cons} \\ +\text{back} \end{bmatrix} \rightarrow [-\text{back}] \;/\; \begin{bmatrix} -\text{cons} \\ -\text{back} \end{bmatrix} C_0 \;\_\_\_$$

The *rule* in (9) is interpreted as follows. The *focus* [−consonantal, +back] to the left of the arrow defines the input to the alternation as the back vowels. The matrix [−back] to the right of the arrow indicates the feature change introduced by the rule – the *structural change* (SC). The slant / is read "in the context." The accompanying *environment dash* ___ locates the focus relative to the conditioning context. Finally, C_0 denotes a string of zero or more consonants. The statement in (9) thus says that back vowels are changed to the corresponding front vowels when preceded by a front vowel with an intervening string of zero or more consonants. The expression formed by combining the rule's focus and conditioning context (if any) is known as the *structural description* (SD) *of the rule*. Thus, A → B / ___ C has the SD AC; the SD of A → B / C ___ is CA; and the context-free rule A → B has A as its SD.

The vowel fronting in (9) illustrates the most common type of sound change: a sound *assimilates* a feature of the local environment. Our feature system can characterize assimilation by the extent to which the feature specifications in the structural change match the specifications in the conditioning context. Chamorro exhibits assimilation for backness. Since assimilation is the most common type of sound change, we can begin to understand how the class of possible phonological changes [x] → [y] might be constrained. As a first approximation, we expect [x] to become [y] in those contexts [z] where [z] has one or more of the features mentioned in [y].

Finally, with features we can begin to give a plausible answer to the question posed in the Introduction about sound change. Recall that the problem is to explain how it is possible to systematically modify our speech in a very precise but largely unconscious way as we move from one dialect (regionally or socially defined) to another. We may now say that the information the brain sends to the vocal apparatus in order to articulate a given lexical item is represented in memory in the form of a string of speech sounds, where each sound is a distinctive feature matrix. For example, in many Southern dialects of American English, the vowel [ɛ] is raised to [ɪ] before a nasal consonant so that *pen* becomes homophonous with *pin* and *hem* with *him*. At the level of detail we are considering here, the word *pen* will have the representation in (10).

(10) $$\underset{[\mathrm{p}]}{\begin{bmatrix} +\text{cons} \\ -\text{continuant} \\ +\text{labial} \\ -\text{voiced} \end{bmatrix}} \quad \underset{[\varepsilon]}{\begin{bmatrix} -\text{cons} \\ -\text{high} \\ -\text{low} \\ -\text{back} \\ -\text{ATR} \end{bmatrix}} \quad \underset{[\mathrm{n}]}{\begin{bmatrix} +\text{cons} \\ +\text{nasal} \\ +\text{coronal} \end{bmatrix}}$$

As a first approximation, we may say that this dialect has the rule stated in (11).

(11) $$\begin{bmatrix} -\text{cons} \\ -\text{low} \\ -\text{back} \\ -\text{ATR} \end{bmatrix} \rightarrow [+\text{high}] \;/\; \_\_\_\_ \; [+\text{nasal}]$$

Application of this rule thus alters the realization of every [ɛ] before a nasal – regardless of the lexical item in which it occurs. Speakers applying this rule in their speech thus compute a different pronunciation for *hem* and *pen*, one that matches *him* and *pin*.

Assuming that speech sounds are bundles of distinctive features, we conceive of sound change as an alteration of the plus/minus specifications of the entries in the feature matrix. Let us examine some additional vocalic alternations to begin learning the kinds of changes these segments are subject to. The Bizcayan dialect of Basque (de Rijk 1970) has the five-vowel system [i,e,a,o,u]. The indefinite morpheme *bat* 'one' and the definite suffix show an alternation between [a] and [e].

(12)

noun	indefinite	definite	
sagar	sagar bat	sagar-a	'apple'
gison	gisom bat	gison-a	'man'
buzten	buztem bat	buzten-a	'tail'
belaun	belaum bet	belaun-e	'knee'
cakur	cakur bet	cakur-e	'dog'
agin	agim bet	agin-e	'tooth'
mutil	mutil bet	mutil-e	'boy'

Examination of the data reveals that these morphemes are pronounced with an [e] when a high vowel [i,u] occurs in the immediately preceding syllable and with an [a] when the mid [e,o] or the low [a] precedes. Since the indefinite morpheme has the shape *bat* in isolation, the *bet* alternant appears to be the derivative one. We can characterize the relationship between the alternants by the rule in (13). It says that a low vowel changes its values for [low] and [back] to minus when preceded by a high vowel. (This rule may be interpreted as a restricted assimilation of [−low] from a subset of the nonlow vowels – those that are [+high]. The change to [−back] is best seen as a subsidiary adjustment.)

(13) $$\begin{bmatrix} -\text{cons} \\ +\text{low} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{low} \\ -\text{back} \end{bmatrix} \;/\; \begin{bmatrix} -\text{cons} \\ +\text{high} \\ -\text{low} \end{bmatrix} C_0 \_\_\_$$

The rule thus explains why the definite and indefinite suffixes change their pronunciation after *cakur* 'dog', *belaun* 'knee', *agin* 'tooth', and *mutil* 'boy' but not after the other nouns in (12). We see that the high vowels [i] and [u] form a natural class in triggering this rule. Given our feature system, the simple designation [+high] can isolate this class of vowels. [+low] picks out the [a] vowel as the rule's focus – the only segment that alternates in this context in Basque. It shows that a single segment may constitute a natural class.

The high vowels figure in another rule of this Basque dialect that operates in the definite forms of the nouns in (14). These stems end in a high vowel. As predicted by rule (13), the low vowel of the suffixes has changed to [e].

(14)

noun	indefinite	definite	
erri	erri bet	erriye	'village'
ari	ari bet	ariye	'thread'
buru	buru bet	buruwe	'head'
iku	iku bet	ikuwe	'fig'

The *semivowels* or *glides* [y] and [w] are close kin to the corresponding high vowels [i] and [u]. If one articulates an [i] and then slowly constricts the tongue body, the pronunciation shades into the *jod* [y]. [u] shades into [w] under similar constriction. Intuitively, [y] and [w] are consonantal variants of the vowels [i] and [u]. The emergence of the semivowels in the definite forms of (14) can be explained in the following way. Vowel sequences (V+V) are phonologically unstable; affixation of the definite suffix [-a] creates such a *hiatus*. One response is to separate the vowels by a "dummy" or "generic" consonant: V+V → VCV. Basque realizes this dummy as the consonantal variant of the preceding high vowel. In an intuitive sense (which will be made more precise in chapter 4), the features of the preceding high vowel seep into this position, giving rise to [y] and [w]. This idea is sketched in (15).

(15) $$\begin{array}{ccc} V & C & V \\ | & & \\ \begin{bmatrix} +\text{high} \\ \pm\text{back} \end{bmatrix} & & \end{array}$$

1.5 Round and Central Vowels

We have characterized vowel quality in terms of tongue body position. The configuration of the lips may also play a distinctive role. Vowels with essentially the same tongue positioning may nevertheless contrast in quality. For example, the vowels in German *Tier* 'animal' and *Tür* 'door' are both articulated with the tongue in a relatively high front position. The difference lies in the fact that for the vowel of *Tür* the lips are rounded and/or protruded while the vowel of *Tier* is articulated with the lips in a neutral or spread position. To account for this additional phonetic dimension, we augment the feature system with [±round]. In German each of the nonround front vowels [i,e,ɛ] has a distinctive rounded counterpart. We follow German orthography and transcribe the front rounded vowels with the umlaut as [ü,ö,ɔ̈]. Other languages with one or more front rounded vowel include French, Swedish, and Finnish. In similar fashion, many languages (e.g., Russian, Mandarin, Korean) have one or more back unrounded vowel. As with the front rounded vowels, the symbolization of the back unrounded vowels is varied. (16) displays the two most common systems of vocalic transcription: the system of the International Phonetic Association (IPA) and a system more commonly employed in America. The IPA tends to utilize special symbols while the American system relies more on diacritics. In generative phonology, the American system has been more commonly employed. One reason is that it more closely reflects the idea that sounds are composed of features; it also has the advantage of having fewer basic symbols to learn and is easier to deploy typographically (at least until the advent of word-processing software, which brings the special IPA fonts within easy reach). In general, the American system will be employed here, except that IPA [ɯ] and [ɤ] will be used instead of [ï] and [ë].

(16)

	[−back]		[+back]	
	[−round]	[+round]	[−round]	[+round]
[+ATR]				
[+high]	i	y	ɯ	u
[−high]	e	ø	ɤ	o
[−ATR]				
[+high]	ɩ	Y		ɷ
[−high]	ɛ	œ	ʌ	ɔ

IPA transcription of front and back round and nonround vowels

	[−back]		[+back]	
	[−round]	[+round]	[−round]	[+round]
[+ATR]				
[+high]	i	ü	ï	u
[−high]	e	ö	ë	o
[−ATR]				
[+high]	I	Ü	Ï	U
[−high]	ɛ	ɔ̈	ʌ	ɔ

Alternative "American" transcription of front and back round and nonround vowels

Complicating the matter of transcription further is the fact that many philological and linguistic traditions specializing in particular language families have developed their own transcription systems (some of which take into account the graphic traditions of the individual languages). Thus, in Slavic linguistics the high back unrounded vowel is traditionally transcribed as [y] and the front palatal glide jod as [j]. In general, generative phonologists prefer to retain the transcription of their sources; and so the confusing plethora of symbols is a fact of life that readers of the literature must make the best of. To minimize confusion, IPA or "American" equivalents are normally indicated. Pullum and Ladusaw 1986 is a very useful guide to phonetic symbolization.

To show the independence of lip rounding and tongue position, some minimal pairs are cited from German (17a) and Korean (17b).

(17) a.	Tier	[i]	'animal'	Tür	[ü]	'door'
	Sehne	[e]	'tendon'	Söhne	[ö]	'sons'
b.	kul	[u]	'cave'	kɯl	[ɯ]	'script'
	sol	[o]	'brush'	sɤl	[ɤ]	'frost'

Turkish has the vowel system shown in (18a), with matched sets of round and nonround vowels in the high front and back regions. For many Turkish suffixes, the feature values for [back] and [round] are determined by the corresponding values of the root vowels and thus do not form lexical contrasts. The data in (18b) illustrate this *vowel harmony*. The plural suffix takes the back vowel alternant *-lar* when the root contains one of the back vowels [u,ɯ,o,a] and the front vowel alternant *-ler* when the root contains one of the front vowels [i,ü,e,ö]. Similarly, the [±back] value of the accusative suffix is determined by the root; it is [+back] when the root contains a back vowel and [−back] when the root contains a front vowel.

(18) a.

i ü	ɯ u
e ö	a o

	i	e	ü	ö	ɯ	u	o	a
high	+	−	+	−	+	+	−	−
back	−	−	−	−	+	+	+	+
round	−	−	+	+	−	+	+	−

b.

noun	pl.	acc.	
dal	dal-lar	dal-ɯ	'branch'
kol	kol-lar	kol-u	'arm'
kɯz	kɯz-lar	kɯz-ɯ	'daughter'
kul	kul-lar	kul-u	'slave'
yel	yel-ler	yel-i	'wind'
göl	göl-ler	göl-ü	'sea'
diš	diš-ler	diš-i	'tooth'
gül	gül-ler	gül-ü	'rose'

If we let the variable α range over plus and minus, then the backness harmony rule can be expressed as (19a). This rule says that for any given vowel V (= [−consonantal]), its value for [back] is the same as the [back] value of the preceding vowel; C_0 stands for zero or more consonants and is a notational device allowing the rule to apply regardless of the number of consonants intervening between the vowels.

(19) a. $V \rightarrow [\alpha\text{back}] \,/\, \underset{[\alpha\text{back}]}{V} \; C_0 \_\_\_$

b. $\underset{[+\text{high}]}{V} \rightarrow [\alpha\text{round}] \,/\, \underset{[\alpha\text{round}]}{V} \; C_0 \_\_\_$

Suffixal high vowels, such as the accusative, also harmonize in rounding with the root vowel in Turkish. Thus, after the [+round] root vowels [u] and [o] we find suffixal [u] if the root vowel is also [+back] and suffixal [ü] if the root vowel is [−back] [ü] or [ö]. We express the rounding harmony with the help of rule (19b).

Within the class of low vowels, different dialects of English (as well as other languages) make various distinctions. Most dialects distinguish the [−back] [æ] of *cat* from the [+back] [ɑ] of *father.* A number of dialects have another vowel [a] intermediate in quality: for example, the Eastern Massachusetts [a] of *car, park, yard,* and so on, and the Chicago vowel of *pop* [a] distinct from the [ɑ] of *balm* and the [æ] of *cat*. Other dialects distinguish the vowel of *bomb* from the [−round] [ɑ] of *balm* and the [−low] [ɔ] of *bought*. This vowel is transcribed as [ɒ] and has the featural representation [+low, +back, +round].

Traditional phonetics recognizes a distinct class of central vowels in the mid and high range as well. Both round and nonround variants are allowed. In the high range they are transcribed as barred [ɨ] and [ʉ]. In the mid range the central vowels include schwa [ə] and the retroflexed (*r*-colored) vowel of *bird* [ɜ] (formed by curling the tip of the tongue back). Most generative phonologists have eschewed introducing a distinct central category in addition to front and back; they have attempted to accommodate these vowels in other ways. In many languages the central vowels arise as reduced variants of front and back vowels in unstressed position and so do not contrast with full vowels: in English, for example, compare the high reduced vowel [ɨ] in the initial syllable of *demonic* with the mid schwa [ə] of *atomic*. In other cases the feature [round] may be invoked. For example, the Togolese language Lama (Ourso 1989) exhibits a three-way contrast among the high vowels: [−back, −round] [i] (e.g., *litə* 'tease'), [+back, +round] [u] (e.g., *lutə* 'stir up'), and a distinct vowel intermediate in quality between these two (which Ourso transcribes as schwa but shows to be [+high] on both acoustic phonetic and phonological grounds): *lətə* 'skin an animal'. Since this vowel lacks lip rounding, it could be described as [+back, −round] [ɯ]. But it is unclear whether there is any motivation for grouping it with the [+back] [u] other than the desire to avoid central vowels. Finally, there are cases in which this maneuver is precluded: Clements (1991a) mentions a dialect of Swedish spoken in Finland that distinguishes among three high vowels all of which are [+round]: the front [ü] of *dyr* 'expensive', the central [ʉ] of *bur* 'cage', and the back [u] of *bor* 'lives'. This is a case where a third category must be introduced (see section 9.3). This discussion can be summarized with the table in (20), where the central vowels are distinguished by the lack of a specification for [back].

(20)

	æ	a	ɑ	ɒ	ə	ɨ	ʉ
high	−	−	−	−	−	+	+
low	+	+	+	+	−	−	−
back	−		+	+			
round	−	−	−	+	−	−	+

1.6 Features for Consonants

Table (1) of the most frequent consonants is repeated in (21).

(21)

	labial	coronal	velar
voiceless stops	p	t	k
voiced stops	b	d	g
voiceless fricatives	f	s	x
voiced fricatives	v	z	ɣ
nasals	m	n	ŋ

Corresponding to the phonetic dimensions discussed earlier, we introduce the features [±continuant], [±voiced], and [±nasal]. The stops (produced with a constriction blocking airflow) are [−continuant], while the fricatives (which permit turbulent airflow) are [+continuant]. Nasals (produced with lowered velum) are [+nasal], while oral segments are [−nasal]. [+voiced] denotes sounds produced with a vibration of the vocal folds; [−voiced] sounds lack this vibration. The table in (22) illustrates how these features differentiate the coronal consonants of (21). The labials and velars receive a corresponding analysis.

(22)

	t	d	s	z	n
continuant	−	−	+	+	−
voiced	−	+	−	+	+
nasal	−	−	−	−	+

The [continuant] value of the nasal consonants is not completely clear. Since airflow is not interrupted, nasals can be prolonged like fricatives (humming) and consequently might be considered [+continuant]. On the other hand, they are produced with an oral closure and hence might better be treated as [−continuant]. The latter decision is supported by the frequent shift of fricatives to stops after a nasal, a change that can be viewed as the assimilation of [−continuant] from the nasal (compare Spanish *la* [ɣ]*ata* 'the cat' fem. but *u*[ŋ] [g]*ato* 'a cat' masc.).

1.6.1 Place of Articulation: The Articulator Theory

Traditional phonetics recognizes some eleven distinct points of articulation dispersed along the vocal tract, as depicted in figure 1.1. Bilabials are produced by a constriction at the lips, while for labiodentals the constriction is formed by the lower lip and the upper teeth. The (inter)dentals, dental-alveolars, and alveopalatals constrict the tongue blade respectively at the back of the upper teeth, the alveolar ridge, and the roof of the mouth at the point where it slants upward toward the soft palate. Velars and postvelars (uvulars) form constrictions with the tongue dorsum. In pharyngeals the tongue root approximates the back wall of the pharynx, while the laryngeals utilize the vocal folds as articulators.

Recent research amalgamates some of the traditional place-of-articulation categories according to the active articulator that forms the consonantal constriction: the lower lip [labial], the tongue blade [coronal], the tongue body [dorsal], the tongue root [radical], and the vocal folds [laryngeal]. The oral articulators [labial],

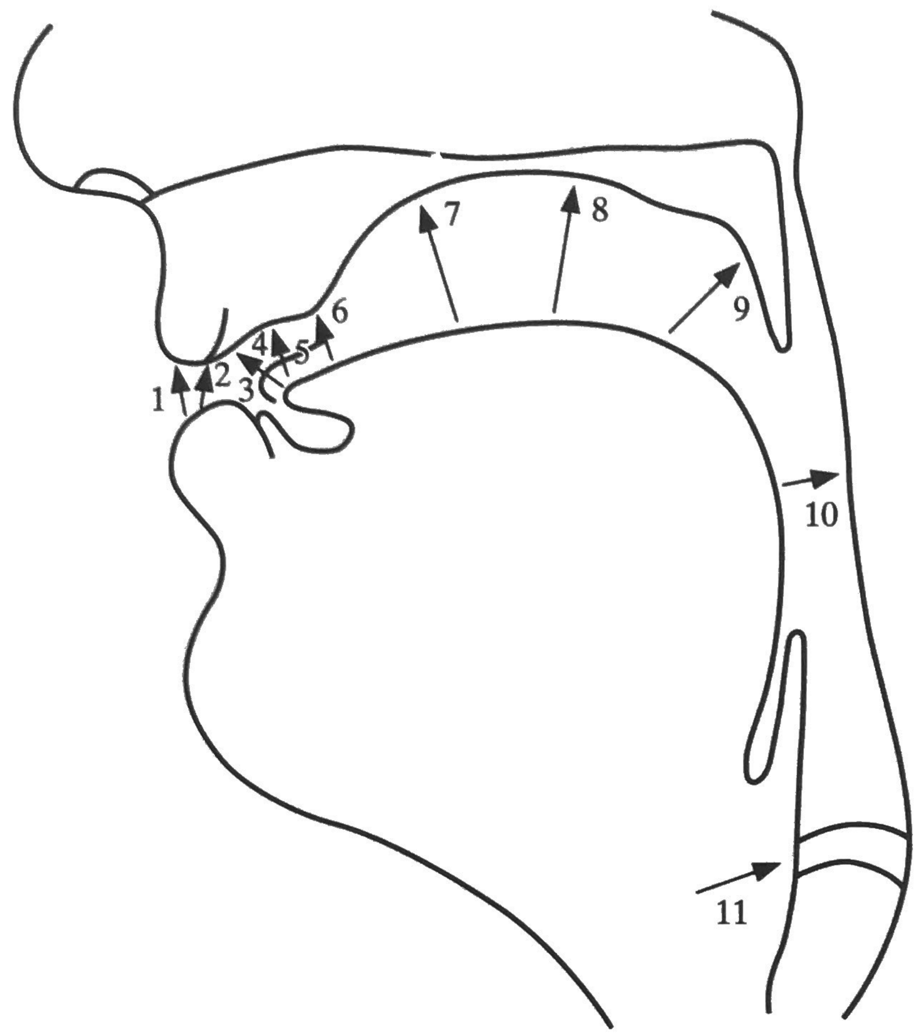

Figure 1.1 Places of articulation: 1 bilabial; 2 labiodental; 3 dental; 4 alveolar; 5 retroflex; 6 alveopalatal; 7 palatal; 8 velar; 9 uvular; 10 pharyngeal; 11 laryngeal. (Adapted from Ladefoged 1975:6.)

[coronal], and [dorsal] function as a group for one of the most widespread phonological processes: nasal assimilation to the place of articulation of a following consonant. This process is reflected in the varying realizations of the prefix in *co*[m]-*pact*, *co*[n]-*tact*, *co*[ŋ]-*gress*. It suggests that there is a phonological dimension "place" that has more than two values, including at least labial, coronal, and dorsal. In chapter 4 we will undertake a more precise characterization of the place features. For now, we content ourselves with [±labial], [±coronal], and [±dorsal].

This interpretation of place of articulation departs from earlier treatments in generative grammar where labial, dental-alveolar, palatal, and velar were differentiated by two binary cuts. The precise subgroupings varied somewhat from one researcher to another. As with so many other aspects of generative theory, the interpretation put forth in Chomsky and Halle's (1968) landmark study *The Sound Pattern of English* (henceforth *SPE*) served as the unofficial standard. In *SPE* the features [anterior] and [coronal] partitioned the oral place of articulation as in (23).

(23)

	anterior	coronal
labial	+	−
dental-alveolar	+	+
alveopalatal	−	+
velar	−	−

The feature [coronal] was defined essentially as here, while [anterior] split the oral cavity into two broad regions at the point where the alveolar ridge rises upward (roughly where the first consonant in *ship* is articulated). However, subsequent research has shown that [coronal] and [anterior] do not have the same status. While there are many sound changes that distinguish dentals and palatals from labials and velars, few convincing cases partition the four major places of the oral cavity as {labial, dental} vs. {palatal, velar}, as predicted by the [±anterior] distinction proposed in (23). Furthermore, while changes can be found that relate velars and labials on the one hand and dentals and palatals on the other, these changes never occur in the same context – a surprising empirical gap if [±anterior] is defined as claimed in (23).

1.6.2 Subsidiary Place Distinctions

Each of the articulators imposes further refinements that have traditionally been subsumed under the label "place of articulation." In this section we will briefly survey these subcategorizations.

Labial

Labial consonants can be distinguished as *bilabials* and *labiodentals*. In the former, the consonantal constriction is made by the two lips; the latter are articulated by an approximation of the lower lip and the upper teeth. The customary symbols are shown in (24).

(24)	bilabial	labiodental
stop	p b	
fricative	Φ β	f v
nasal	m	ɱ

The fricatives in [f]*ile* and [v]*ile* are labiodental; the fricative of Spanish *la* [β]*arca* 'the boat' is bilabial. The nasal [ɱ] appears in some pronunciations of *emphasis*. It invariably arises from a rule of place assimilation and is not an independently occurring segment. A few languages contrast the bilabial and labiodental fricatives. Ladefoged (1975) cites such minimal pairs as *èβè* 'Ewe' vs. *èvè* 'two' and *éΦá* 'he polished' vs. *éfá* 'he was cold' from the West African language Ewe. For the fricatives, languages more commonly select the labiodentals [f,v] than the bilabials [Φ,β]. This asymmetry is customarily explained on acoustic grounds. The labiodentals have a higher degree of turbulence and hence are more salient than the mellow-sounding bilabials. Some researchers have distinguished the two subclasses by invoking a feature [strident]: the harsher labiodentals are [+strident] in opposition to the [−strident] bilabials. But since stridency is a property of fricatives and not nasals, this feature cannot be the one that is assimilated in *emphasis*. If this type of assimilation is phonological and not merely phonetic *coarticulation* (overlap of articulatory gestures), then an additional feature must be introduced to group [ɱ] with [f] and [v]. We leave this question open, pending further evidence.

Coronal

The coronal articulator exhibits the largest number of refinements. *Dental-alveolar* versus *alveopalatal* is the most basic distinction. In the former class the front portion of the tongue creates a constriction either at the back of the upper teeth (pure dentals) or on the hard gum ridge immediately behind (alveolars). The alveopalatals (a term often abbreviated to palatals) are produced with a constriction farther back, at the point where the roof of the mouth begins to slant upward toward the soft palate. Many languages draw coronals from these two regions. Minimal pairs from English are cited in (25). They are distinguished with the help of the *SPE* feature [anterior] but its use is explicitly restricted to the coronals.

(25)	[+anterior]	[−anterior]				
	s	š	(IPA ʃ)	sun	vs.	shun
	z	ž	(IPA ʒ)	miser		measure
	t	č	(IPA tʃ)	tin		chin
	d	ǰ		dump		jump
	n	ñ		money		onion

The palatal nasal is found in some pronunciations of *onion* (i.e., *o*[ñ]*on*) and occurs in Spanish.

The dental-alveolars are more common than the alveopalatals. Virtually every language selects at least one (and usually several) consonants from the [+anterior] column while many shun [−anterior] altogether. This asymmetry is also reflected in languages like Spanish where the [−anterior] [ñ] and [ʎ] are replaced by the corresponding [+anterior] [n] and [l] in the coda of a syllable: compare *Do*[ñ]*a* 'lady' but *Do*[n] and *Donce*[ʎ]*a* 'lass' but *Donce*[l] 'lad'. A neutralization of the opposition from [+anterior] to [−anterior] is much less common in this position.

The coronals are partitioned in two other ways as well. The traditional phonetic category of *interdentals* is comprised of the voiceless and voiced fricatives represented by the initial consonants of English *thin* [θɪn] and *then* [ðɛn]. They are produced by placing the tongue blade parallel to the roof of the mouth in such a way that the tip touches or projects slightly beyond the upper teeth. Air is expelled over the tongue, between the teeth. In English these fricatives are opposed to the [s] of *sin* and the [z] of *Zen* in which the tip of the tongue makes a constriction at the alveolar ridge. Many languages of Australia distinguish between dental and alveolar stops. Also, in the Irish English brogue the interdentals are replaced by dental stops, creating a dental-alveolar contrast (marked by the presence versus absence of the underbridge) between *thin* [t̪]/*though* [d̪] and *tin* [t]/*dough* [d] (Wells 1982:428). Most researchers have sought to distinguish the interdental, dental, and alveolar categories by differences in tongue shape rather than as differences in point of articulation, adding precision to the traditional phonetic distinction of *apical* (tongue-tip) versus *laminal* (tongue-blade) articulation. In *SPE* the apical-laminal distinction is implemented by a feature [distributed]: [+distributed] are sounds produced with a constriction that extends a considerable distance parallel to the direction of airflow; [−distributed] sounds are produced with a constriction that extends for just a short distance along the direction of airflow. On these grounds, the interdentals are [+distributed]. The [±distributed] feature can also be used to distinguish the retroflex consonants found in many languages of India; these consonants are produced by curling the tongue back to make a constriction with the tip or the underside of the tongue in the alveopalatal region. The [t] in [ṭ]*rip* is retroflexed for many English speakers. Retroflexion is commonly denoted by underdotting. The IPA uses an undertail. The table in (26) summarizes the partitioning of the coronal articulator induced by the features [anterior] and [distributed].

(26)

		[anterior]	[distributed]
interdental	θ	+	+
dental-alveolar	s	+	–
alveopalatal	š	–	+
retroflex	ṣ, ʂ	–	–

Traditional phonetics distinguishes another point of articulation at the flat upper surface of the hard palate between alveopalatals and velars. It is usually referred to as (pure) *palatal;* it is articulated with the tongue blade and hence a species of coronal. Sounds in this category include the voiceless and voiced palatal stops of Hungarian *kutya* 'dog' and *Magyar* 'Hungarian' (transcribed [c] and [ɟ] by the IPA) and, according to Ladefoged (1975), the German ich-laut (a voiceless fricative) in words such as *ich* 'I' and *nicht* 'nothing', transcribed with the [ç]. Rather than adding another point-of-articulation distinction within the coronals, most phonologists have tried to accommodate these segments through appeal to the stricture features of [continuant] and [strident] (see section 1.7).

Affricates

Languages frequently distinguish one or more *affricates* among their coronal consonants. For example, Russian opposes its dental stop [t] and fricative [s] to the affricate [t^s]: [t']*elo* 'body', [s']*elo* 'village', [t^s]*eloe* 'entire' neut. sg. (the apostrophe in Russian [t'] and [s'] denotes palatalization; see section 1.10). In German the stop [t] of *Teller* 'plate' is opposed to the affricate [t^s] that begins *Zelle* 'cell'. The affricate is produced with an initial closure; but unlike the closure of a stop, the closure of an affricate is released gradually, sounding very much like a fricative. The affricates cannot be treated as a stop-fricative sequence since they freely occur in so-called CV languages that otherwise do not allow consonant clusters. Also, some languages such as Polish distinguish affricates from the corresponding consonantal sequence; compare the cluster of *trzy* [tšɨ] 'three' with the affricate in *czy* [$t^š$ɨ] 'whether'. The interpretation of affricates is controversial. While most other sounds combine their features simultaneously, the affricate seems to require an internal sequencing of the feature [continuant].

(27)

```
      [+coron]                  [+coron]
         |                         |
      [+anter]                  [-anter]
       /    \                    /    \
[-contin][+contin]        [-contin][+contin]
       [tˢ]                      [tˇˢ]
```

This representation is supported by phonological rules that treat the left face as a stop and the right as a fricative. In the alveopalatal-palatal region, the affricates [$t^{\check{s}}$,$d^{\check{z}}$] are preferred over the stops [c,ɟ]. Thus, when [t] and [s] assimilate the palatality of a front glide [y], the stop [t] typically changes to the affricate [$t^{\check{s}}$] while [s] remains a fricative and appears as [+continuant] [š]: *met you* → *me*[$t^{\check{s}}$]*you* but *miss you* → *mi*[š]*ou*. The affricates are tabulated in (28) in voiceless-voiced pairs, along with the most common transcriptional alternatives.

(28)

	[+anterior]	[−anterior]
stop	t, d	c, ɟ
fricative	s, z	š, ʒ
		ʒ, $d^{\check{z}}$
affricate	t^{s}, d^{z}	$t^{\check{s}}$, ǰ
	ȼ	č, ǯ
		tʃ,

Dorsal Consonants

The tongue body is capable of producing constrictions over a broad spectrum of the vocal tract, as shown by points 8 and 9 in figure 1.1. *Prevelars* and *velars* require essentially the same tongue body positioning as the front and back vowels, respectively, and so can be distinguished by the feature [±back]. We can see this in the varying realizations of [k] in response to the front versus back vocalic environment: compare the relatively front prevelar of the initial stop in *keep* with the relatively back velar of *coop*.

In *uvulars*, the tongue dorsum makes a constriction at the uvula. Arabic opposes voiceless stops at the velar and uvular points of articulation: velar *kalb* 'dog', uvular *qalb* 'heart'. The Parisian [ʁ] of *rouge* 'red' is a voiced uvular approximant. Aside from *k* (velar) and *q* (uvular), the transcription of the prevelar, velar, and uvular varieties of dorsal consonants is not standardized. Here a prime will be used for prevelars and underdotting for uvulars. IPA alternants for the uvulars are also shown.

(29)

	prevelar	velar	uvular	
voiceless stop	k′	k	q	
voiced stop	g′	g	g̣	[ɢ]
voiceless fricative	x′	x	x̣	[χ]
voiced fricative	ɣ′	ɣ	ɣ̣	[ʁ]
nasal	ŋ′	ŋ	ŋ̣	[ɴ]

Gutturals

Traditional grammar groups laryngeal, pharyngeal, and sometimes some uvular consonants under the category of *gutturals*. Contrasts among these three points of articulation are found in Classical Arabic. *Pharyngeals* are formed by a constriction at the back wall of the pharynx with the root of the tongue; [ħ,ʕ] and [ḥ,9] are transcriptional alternatives. *Laryngeals* have the vocal folds as articulators: the glottal stop [ʔ] constricts the folds to briefly shut off airflow while the [h] permits air to pass through the glottis. Some minimal pairs are cited in (30b).

(30) a.

	uvular	pharyngeal
voiceless	ẋ	ḥ
voiced	ɣ̣	ʕ

laryngeal
h
ʔ

b.

ẋaali	'my maternal uncle'
ɣ̣aali	'expensive'
ḥaali	'my condition'
ʕaali	'high'
haal	'mirage'
ʔaal	'family, kin'

McCarthy (1991) documents the phonological relevance of the gutturals. Let us look briefly at two types of sound change that single out these consonants. The gutturals frequently condition lowering in an adjacent vowel. For example, the root vowel in Classical Arabic verbs undergoes *ablaut* (phonetically unconditioned change of a root vowel) in forming the perfect (C*a*CVC) versus imperfect (CCVC) distinction. The high vowels [i] and [u] ablaut with [a]: *katab – ktub* 'write', *ḥamal – ḥmil* 'carry'. However, roots whose second or third consonant is a guttural show the low vowel in both the perfect and imperfect: *saʔal – sʔal* 'ask', *ðahab – ðhab* 'go', *fataḥ – ftaḥ* 'open', *baʕat – bʕat* 'send'. Many Bedouin Arabic dialects shun preconsonantal gutturals, metathesizing the preceding vowel (VGC → GVC) to avoid closing the syllable with a guttural. This process distorts the underlying morphological templates in the following paradigms from the Hijaazi dialect of Saudi Arabia.

(31)

template	plain root	guttural root
CaCC+a	sawda 'black'	bɣ̣aθa 'grey' dhama 'dark red'
ma+CCuuC	maktuub 'written'	mʕazuum 'invited' mḥazuum 'tied' mẋaṣuur 'neglected'

Having surveyed the various points of articulation, we might ask what evidence justifies grouping the consonants in terms of the three major articulators: labial, coronal, dorsal? After all, in our survey we have examined some ten different categories dispersed along the vocal tract. Traditional phonetics sees these as ten separate points along the dimension of consonantal place. Hybrid labels like labiodental might suggest placing [f] between the bilabial and dental categories, sharing the lower lip of [p] and the upper teeth of [t] and [θ]. Yet we have assigned [f] and [v] the labial articulator, implying that they have more in common with [p] than with [t]. Several pieces of evidence support this interpretation. Many phonological systems select a stop and corresponding fricative from each of the oral place regions. For example, typically [k] is paired with [x], [t] with [s], and

[p] with [f]. A system where [t] pairs with [f] (instead of [s] or [θ]) and [p] with [Φ] is unheard of. The claim that [f] and [v] are basically labial is also reflected in sound change. For example, many languages spirantize stops to fricatives (lenition), so that the velar [k] is replaced with [x]. In these languages we find [p] changing to [f] – not to [θ] or [s]. This process is often inverted (fortition); once again, [f] typically alternates with [p], not with [t]. This point is also explained if it is the feature [continuant] that changes while the [labial] articulator remains constant.

1.7 Stricture

Traditional phonetics recognizes three categories of consonantal constriction: *stops* (also called *occlusives*) interrupt airflow, *fricatives* constrict airflow in such a way as to generate turbulence, and *approximants* have nonturbulent airflow. Common representatives of each category are depicted in (32).

(32)	occlusives	p	t	č	k
	fricatives	f	s	š	x
	approximants	w	l,r	y	

Occlusives are distinguished from fricatives by the feature [continuant]. [−continuant] sounds are produced with an oral stricture that interrupts airflow. On these grounds, [p,t,č,k] are [−continuant] while [f,s,š,x] are [+continuant]. Nasals [m,n,ñ,ŋ] fail to participate in a stop-fricative opposition the way [p] and [f] do, but they induce occlusion on following consonants and for this reason are usually classed as [−continuant]. Approximants include the liquids [l] and [r] and the glides [w] and [y]. These sounds form a natural class in virtue of their distribution in syllable structure rather than because of any property they impart to neighboring sounds via assimilation. For example, in English syllable onsets may include a stop plus liquid or glide (*clean* [kl], *cream* [kr], *queen* [kw], *cue* [ky]) but not a stop plus nasal such as [kn] (in *knee* the *k* is silent). While the approximants could be defined indirectly as nonnasal sonorants, some researchers have preferred the feature [±approximant] so that the shared distribution of the liquids and glides can be expressed directly.

Having surveyed the major stricture categories, we now consider each of them in turn. Most phonological systems select an occlusive for each oral articulator, the coronal often taking two – typically a [+anterior] such as [t] and a [−anterior] such as [č]. Affricates are more common than stops at the alveopalatal point of articulation; they can also be found at the lips and dorsum (e.g., standard German [pf] and dialectal [kx]). Because of the difficulty of interrupting airflow at certain points of articulation (e.g., interdental and pharyngeal), there are more continuants than stops. In some systems, the interdentals [θ,ð] are distinguished from the harsher [s,š], the latter termed *sibilants*. Some phonologists have proposed a feature [+strident] (greater turbulence) to distinguish the sibilants from the [−strident] interdentals. Among the approximants, the liquids [l,r] and the glides [w,y] form easily distinguished subsets. In Maddieson's (1984) survey of more than three hundred phonological inventories, over 95 percent of languages chose at least one liquid and most (77 percent) more than one, with the lateral being slightly more popular. Liquids are prone to dissimilate; in Latin [l] turned to [r] when the stem contained a lateral. This dissimilation is evident in the distribution of the *-al* and *-ar* alternants of the adjectival suffix in *pap-al, tot-al, coron-al, guttur-al* but *vel-ar, pol-ar, angul-ar*. A *lateral* is produced by making a constriction with the central portion of the tongue but lowering one or both margins so that air flows out the side of the mouth. Laterals may be aspirated or voiceless (transcribed as a barred [ɬ]); they can be affricated; and while typically coronal (with the dental-alveolar [+anterior] [l] favored over the alveopalatal [−anterior] [ʎ]), they can be articulated with the tongue body. Whether coronal versus velar lateral is a possible phonemic contrast is unclear. *Rhotics* come in several varieties as well. The trilled versus flap opposition of Spanish *perro* 'dog' vs. *pero* 'but' is well known. In a *trill*, the articulator is held loosely so that it can be set in vibration by the passing air. According to Ladefoged (1982:153), "a *flap* is caused by a single contraction of the muscles so that one articulator is thrown against another. It is often just a very rapid articulation of a stop closure," as produced by English speakers who flap the medial consonants in *latter, ladder, tanner*. Phoneticians have devised a panoply of special symbols to distinguish among the various rhotics. For example, the French uvular rhotic is indicated by a small capital, inverted to distinguish the Parisian fricative/approximant [ʁ]*ouge* 'red' from the Midi's trilled [ʀ]*ouge*. The rhotic of English *red* is a weak approximant (IPA [ɹ]) with no contact between the tip of the tongue and the roof of the mouth. It patterns with the glides in allowing the flapping of a following dental stop, which itself is transcribed [ɾ]: *writer* [ayɾ], *outer* [awɾ], *Carter* [aɹɾ] but *welder* [ɛld]. We will distinguish the rhotics and laterals with the help of a feature [±lateral] whose articulatory correlate is the presence or absence of lowering at the tongue margin(s). Finally, the pharyngeals [ḥ,ʕ] and laryngeals [h,ʔ] are sometimes treated as approximants instead of fricatives.

Sound changes from left to right on the stop-fricative-approximant dimension are known as *weakenings* (*lenition*) while changes from right to left are *strengthenings* (*fortition*). Postvocalic context is the most typical environment for the change from stop to fricative (a process also known as *spirantization*). This is the environment where Tiberian Hebrew changes its stops [p,t,k] and [b,d,g] to the fricatives [f,θ,x] and [v,ð,ɣ]. Many systems restrict weakening to contexts in which a vowel follows as well as precedes: for instance, intervocalic flapping of dental stops in many English dialects. It is unclear whether spirantization is properly viewed as assimilation of the open position of the neighboring vowel and hence whether vowels are properly viewed as [+continuant]. Fortitions from fricative to stop tend to occur in the complementary set of contexts: postconsonantal and initial. In some cases only certain consonants call for a following occlusive. For example, in Spanish the voiced stops [b,d,g] and the corresponding fricatives [β,ð,ɣ] are in complementary distribution. The fricatives occur after [r] (*cur*[β]*a* 'curve', *ver*[ð]*e* 'green', *lar*[ɣ]*o* 'long') and stops after a (homorganic; i.e., same point of articulation) nasal: *hom*[b]*re* 'man', *don*[d]*e* 'where', *an*[g]*osto* 'narrow') as well as initially ([b]*ola* 'ball', [d]*uro* 'hard', [g]*ato* 'cat'). After the lateral, [d] appears in place of [ð] (*cal*[d]*o* 'broth') while noncoronals are realized as [+continuant] (e.g., *cal*[β]*o* 'bald', *al*[ɣ]*o* 'something'). The fact that nasals favor a following occlusive justifies categorizing them as [−continuant] even though they themselves typically do not spirantize. The inconsistent behavior of the lateral leaves the [±continuant] status of [l] unresolved. *Gemination* (doubling) is a context that inhibits lenition and promotes fortition. Thus, in Tiberian Hebrew postvocalic stops systematically fail to spirantize when geminate. Similarly, geminated approximants often strengthen to homorganic obstruents: for instance, in Faroese (Anderson 1972) **yy* > *ggj* [ǰ] and **ww* > *gv*.

1.8 Major Class Features

The features [sonorant] and [consonantal] each partition the set of speech sounds into two broad classes, as depicted in (33). Like [continuant], these features are not bound to a particular articulator; instead, they specify phonologically critical degrees of constriction imposed by essentially any articulator.

(33)

	[sonorant]	[consonantal]	
vowels	+	−	a,i,u,o,e, . . .
glides	+	−	y,w
liquids	+	+	l,r
nasals	+	+	m,n,ñ,ŋ
obstruents	−	+	t,d,s

[±sonorant] classifies sounds in terms of the effect their stricture has on the flow of air across the glottis and hence the capacity to induce vibration of the vocal folds. Vocal fold vibration is influenced by several factors; but the most important is airflow. The folds cannot vibrate if no air is passing through the glottis. In order for air to flow, the supralaryngeal pressure must be less than the sublaryngeal. The degree of stricture made during the articulation of a sound may increase the supralaryngeal pressure and hence tend to shut off voicing unless other adjustments are made. Stops and fricatives have a stricture that inhibits spontaneous voicing. The stricture associated with [+sonorant] segments does not disrupt airflow enough to inhibit voicing. Thus, the natural state for sonorants is [+voiced] and for nonsonorants (termed *obstruents*) is [−voiced].

Of all speech sounds, vowels and glides are produced with the least constriction and hence allow the freest flow of air. While nasals are articulated with an oral closure, the nasal cavity is open and hence airflow is not impeded. In the production of a lateral, a closure is made; but the lowering of the tongue margins allows sufficient airflow to maintain voicing. Finally, the various rhotics (retroflexed approximants, trills, flaps) either do not make a sufficiently narrow stricture, or if they do, it is not held long enough to inhibit spontaneous voicing.

From these remarks, it is clear that [+voiced] is the natural laryngeal state for a sonorant. A special laryngeal adjustment is required to inhibit voicing. Thus, while voiceless sonorants occur, they are distinctly dispreferred to voiced ones. In obstruents the opposite state of affairs obtains. The optimal stop or fricative is voiceless. A language typically augments its inventory with a voiced obstruent only after the corresponding voiceless one has been chosen. The special status of [+voiced] in obstruents is shown by the fact that in many languages (e.g., Slavic) obstruents induce voicing assimilation in adjacent obstruents while sonorants are inert, neither causing nor undergoing such changes: compare Russian *pro*[s']-*it* 'requests', *pro*[z']-*ba* 'a request' with *ne*[s]-*ut* 'carries' 3pl., *ne*[s]-*la* past fem. Two additional properties distinguish sonorants. First, in many languages a syllable can only terminate in a sonorant: for example, Lama has words ending in liquids or nasals but none ending in obstruents. Second, in a number of languages (e.g., Lithuanian) a sonorant consonant in the coda of a syllable may count as a tone-bearing unit; obstruents typically shun tones.

The feature [+consonantal] denotes sounds with a radical constriction in the supralaryngeal cavity. Vowels and the corresponding glides are [−consonantal]. Note that in the system being described here the high vowels [i] and [u] and the corresponding glides [y] and [w] have the same feature structure. They differ in terms of their location within the syllable. A vowel occupies the nuclear peak while a glide appears in the margins – the prevocalic onset or the postvocalic coda. The equivalence of [i] – [y] and [u] – [w] is seen in a number of processes. For example, many dialects of Arabic avoid words ending in consonant clusters. When the 2sg. suffix is removed from the stems in *dalw-ak* 'your pail' and *jady-ak* 'your kid', the glides are turned to the corresponding vowels: *dalu* 'pail', *jadi* 'kid'. Another reflection of the close relation between [i] and [y] and between [u] and [w] is seen in Basque. As noted in section 1.4, when the definite suffix *-a* is added to a stem ending in a high vowel, a corresponding glide fills the gap to provide the second syllable with an onset: *ari* 'thread' and *iku* 'fig', but [ari+a] and [iku+a] are realized as [ariya] and [ikuwa] or, depending on the dialect, as [ariye] and [ikuwe] by another rule. When glides occupy the syllable nucleus as part of a *diphthong* (a tautosyllabic sequence of [−consonantal] segments), they are sometimes called *semivowels* and transcribed with an inverted breve ([ai̯], [au̯]). In the early stages of generative phonology, the syllabicity of a sound was characterized by a feature [±syllabic]. This feature stood out from the others, however, in its lack of any precise phonetic correlate. A major theoretical development of the 1970s and 1980s was the articulation of the more traditional conception of the syllable as a prosodic constituent. In current generative thinking, the syllabicity of a sound is a function of its location in this constituent – not a matter of its feature structure. Syllabic consonants are transcribed by an under-ring or understroke: for example, Czech *vlk* 'wolf' is [vl̥k] or [vl̩k].

The table in (34) summarizes the manner-of-articulation features for a representative set of speech sounds.

(34)

	t	s	θ	n	l	r	y	i
consonantal	+	+	+	+	+	+	−	−
sonorant	−	−	−	+	+	+	+	+
approximant	−	−	−	−	+	+	+	+
continuant	−	+	+	−	+	+	+	+
lateral	−	−	−	−	+	−	−	−
nasal	−	−	−	+	−	−	−	−
strident	−	+	−	−	−	−	−	−

1.9 Laryngeal Features

The stops in *s*[p]*in*, *s*[t]*un*, *s*[k]*in* exhibit the optimal laryngeal state for occlusive consonants: voiceless, unaspirated, unglottalized. However, many systems augment their stock of phonemes through modifications of the sound wave introduced by the vocal folds in the larynx. French, Mandarin, and Nootka illustrate the most common options. In French voiceless stops are opposed to voiced, in Mandarin to aspirated, and in Nootka to glottalized (transcribed with a glottal superscript [p^{ʔ}] or apostrophe [p']).

(35) French

[p]as	'not'	[t]u	'you'	[k]uand	'when'
[b]as	'low'	[d]u	'of'	[g]ant	'glove'

Mandarin

pei	'back'	tai	'to bring'	kan	'to do'
p^{h}ei	'to match'	t^{h}ai	'very'	k^{h}an	'to see'

Nootka

pa:-	'go'	ta:-	'long'	kał-	'branch'
p^{ʔ}a-	'give away'	t^{ʔ}aq-	'just'	k^{ʔ}o:-	'a little way'

Many languages combine these options to make a three-way contrast among laryngeal features. For example, in Thai [t] is opposed to [d] as well as to [t^{h}], while in Amharic [t] is opposed to [d] and to [t^{ʔ}]. Finally, some systems such as Hindi make a four-way distinction among laryngeal features, combining both voicing and aspiration (examples from Ladefoged 1975).

(36) Thai

bàa	'shoulder'	dam	'black'		
pàa	'forest'	tam	'to pound'	kàt	'to bite'
p^{h}àa	'to split'	t^{h}am	'to do'	k^{h}àt	'to interrupt'

Amharic

dɨl	'victory'	gərr	'innocent'
tɨl	'worm'	kɨrr	'thread'
t^{ʔ}ɨl	'quarrel'	k^{ʔ}ɨr	'stay away'

Hindi

pal 'take care of'	p^{h}al 'edge of knife'	bal 'hair'	b^{h}al 'forehead'
tan 'mode of singing'	t^{h}an 'roll of cloth'	dan 'charity'	d^{h}an 'paddy'
ṭal 'postpone'	ṭ^{h}al 'place for buying wood'	ḍal 'branch'	ḍ^{h}al 'shield'
kan 'ear'	k^{h}an 'mine'	gan 'song'	g^{h}an 'kind of bundle'

The interpretation of the laryngeal features adopted here is based in part on the model developed by Halle and Stevens (1971). Two phonetic dimensions are distinguished: the amount of space between the vocal folds (glottal width) and the amount of tension in the folds. Given that the voiceless unaspirated unglottalized series defines the optimal laryngeal state, we take it to be the neutral configuration of the glottis. Aspirated sounds are produced by spreading the vocal folds while glottalized sounds constrict the folds. The former are accordingly [+spread gl, −constr gl] while the latter are [+constr gl, −spread gl]. The neutral [p,t,k] are [−spread gl, −constr gl]. The absence of [+constr gl, +spread gl] reflects the impossibility of realizing two opposing mechanical gestures simultaneously. The diagram in (37) depicts this tripartite division of glottal width.

(37)

open				closed
	−constricted +spread	−constricted −spread	+constricted −spread	
	"aspirated"	"plain"	"glottalized"	
	$[p^h]$	[p]	$[p^ʔ]$	

With this background, let us briefly survey the three laryngeal categories. The *aspiration* of a stop as in English $[p^h]$*in* has traditionally been described as a "puff of air" occurring after the release of the stop closure or as a delay in the onset of vocal cord vibration in the production of the following vowel. The Halle-Stevens model sees both of these phenomena as consequences of the instruction to spread the vocal folds. Accordingly, aspirated stops are [+spread gl] and hence [−constr gl]. The voiced aspirates of Hindi are an embarrassment to the traditional "lag in onset of vibration" theory since the folds must be vibrating in the execution of the stop as well as the following vowel. We will follow Halle and Stevens in construing voicing as a function of glottal tension (noting, however, that this interpretation of vocal fold vibration awaits experimental phonetic confirmation). Even though the folds are separated, they may still vibrate if they are held loosely. Thus, [−voiced] will denote sounds produced with greater vocal fold tension, and [+voiced] will denote sounds in which the folds are held more loosely. (Halle and Stevens suggest the features [stiff] and [slack], but this terminology has not been widely accepted. We will retain the traditional label [±voiced] but will follow Halle and Stevens in construing vocal fold vibration as a function of glottal tension.) The table in (38) shows how the stops of French, Thai, and Hindi are characterized in the proposed system.

(38)

	p^h	b^h	p	b
spread gl	+	+	−	−
constr gl	−	−	−	−
voiced	−	+	−	+

Glottalized consonants are produced by constricting the vocal folds. For example, the medial stop in English *button* is glottalized for many speakers: simultaneous with the oral closure, the vocal folds are constricted to yield a [tʔ].

In some casual pronunciations the oral closure may be suppressed, unveiling the laryngeal constriction in the form of a glottal stop: *bu*[ʔ]*on*. In many languages the glottal constriction in stop consonants is enhanced by raising the larynx in the throat. This gesture compresses the air behind the oral closure; when it is released, a sharp, crackling sound is produced. So far as is known, these *ejectives* do not contrast with plain glottalized consonants and so the laryngeal raising, while dramatic, might best be considered a matter of phonetic implementation rather than a distinct phonological category. Such ejectives are found in many American Indian languages, where they are customarily transcribed with an apostrophe [p',t',k'].

A similar phenomenon occurs with the voiced glottalized segments, which are produced by constricting the glottis but slackening the glottal tension to permit vibration. In many cases these consonants are enhanced by a noticeable lowering of the larynx in the throat. This action may decrease the air pressure in the supralaryngeal cavity, leading to an ingression of air when the oral closure is released. Sounds supplemented with this laryngeal lowering are known as *implosives*. They are customarily symbolized as [ɓ,ɗ,ɠ] and are found in many African languages.

Halle and Stevens (1971) suggest dividing the dimension of glottal tension into three categories instead of just two. For the [−spread gl, −constr gl] group, this move allows them to distinguish the voiced stops of Romance languages like French from the initial stops in English *bun, done, gun*. In the former, vocal fold vibration begins more or less simultaneously with the oral closure; in the latter, there is a considerable delay in the onset of voicing until well after the initial closure. Halle and Stevens see the English [b,d,g] as having an intermediate degree of tension between the totally voiceless (stiff) [p,t,k] of *spin, stun,* and *skin* and the fully voiced (slack) [b,d,g] of French. They also exploit the possibility of an intermediate degree of glottal tension to try to explain a puzzling contrast in Korean. As pointed out originally by Kim (1965), Korean stops display three contrasting laryngeal configurations, none of which involve vocal fold vibration: $t^{h}al$ 'mistake', *t*al* 'daughter', *tal* 'moon'. Halle and Stevens suggest placing the [t*] in the aspirated [+spread gl] series but distinguish it from [t^h] as well as from the voiced [d^h] of Hindi by assigning it an intermediate degree of glottal tension. Subsequent research has shown that the "tense" [t*] is actually produced with a constricted glottis but lacks the ejection of the glottalized consonants of the American Indian languages. Iverson (1987) suggests the features [+constr gl, +stiff vf].

The table in (39) summarizes the laryngeal features. If the distinction between the fully voiced [b] of French and the partially voiced [b] of English is simply a matter of different phonetic implementation, we may replace [stiff vf] and [slack vf] with [voiced]. (Voiceless sonorants are transcribed by an underring (e.g., [m̥]) or by small capitals.)

(39)

	spread gl	constr gl	stiff vf	slack vf	(voiced)
p^h	+	−	+	−	−
b^h (Hindi)	+	−	−	+	+

	spread gl	constr gl	stiff vf	slack vf	(voiced)
p	−	−	+	−	−
b (English)	−	−	−	−	+
b (French)	−	−	−	+	+
pˀ	−	+	−	−	−
p* (Korean)	−	+	+	−	−
ɓ	−	+	−	+	+

1.10 Secondary Articulations and Complex Segments

Many languages amplify their segmental inventories by superimposing the vowel quality features of lip rounding and tongue body palatality or velarity on their consonants. Imposition of [+round], [−back], and [+back] produce the *secondary articulations* of labialization, palatalization, and velarization, respectively; pharyngealization arises from retraction of the tongue body and/or tongue root toward the back wall of the pharynx.

Labialization is usually transcribed by a superscripted *w;* a labialized consonant simultaneously combines lip rounding with the primary articulation. Examples of (near) minimal pairs from the West African language Margi (Hoffman 1963) appear in (40).

(40)	pá	'build'	sà	'drink'	gà	'and'
	pʷá	'pour in'	sʷá	'shut'	gʷà	'enter'

Even though rounding is produced by the lips, this gesture is still compatible with a labial constriction; Margi has a full suite of labialized labials: *bʷàŋ* 'hip', *ɓʷà* 'cook', *fʷàŋ* 'hollow', *mʷàl* 'friend', *vʷǐ* 'gourd-plant'. As we might expect, labialized consonants often arise through the assimilation of [+round] from adjacent rounded vowels. For example, in their description of Nootka, Sapir and Swadesh (1939) posit a stem [ki:ɫ] 'making'. When this stem is preceded by the prefix *ʔo-*, the velar is labialized: *ʔo-kʷi:ɫ* 'making it'.

In Russian essentially all consonants are accompanied by a raising of the tongue body as a secondary articulator. If it is [−back], the consonant is *palatalized;* if it is [+back], the consonant is *velarized.* This contrast is fundamental, appearing in prevocalic, preconsonantal, and word-final positions. (In Slavic linguistics, palatalization is usually denoted by the apostrophe; elsewhere, a superscripted *y* is more common. There is no standard notation for velarization; the plain letters are used here with the explicit understanding that the tongue is in a high back position.) The chain of minimal pairs in (41) illustrates this [±back] opposition on Russian consonants.

(41)	m'at'	'to rumple'
	mat'	'mother'
	mat	'checkmate'
	m'at	'rumpled'

Palatalization and velarization freely combine with labials and most dentals in Russian. But in the [−anterior] alveopalatals they are distributed complementarily: [č] is palatalized ([−back]) while, at least in the standard dialect, [š,ž] are velarized ([+back]). This difference shows up in the effects these consonants have on the following vowel: the phoneme /i/ is realized as [−back] [i] after [č] but as [+back] [ɯ] after [š,ž]: *uč'-it* 'teaches', *duš*-[ɯ]*t* 'smothers', *druž*-[ɯ]*t* 'befriends'. Because we have viewed the feature [back] as a dependent of the dorsal articulator, when velar [k] assimilates [−back] it changes its point of articulation to prevelar [k′]; but palatalization of [p] or [t] entails no necessary change in point of articulation. Compare the palatalizations induced by the dat.sg. suffix *-e: ruk-a, ruk'-e* 'hand'; *rabot-a, rabot'-e* 'work', *tolp-a, tolp'-e* 'crowd'. In many cases, however, dental, velar, and sometimes labial consonants change their primary place of articulation to alveopalatal when palatalized by a front vocalic segment, especially the high vowels [i,ü] and the corresponding glides [y,ẅ] ([ẅ] = IPA [ɥ]). Precisely how to express this change is a subject of continued debate in feature theory; see section 9.3 for discussion.

The tongue root or body may be retracted during the production of an oral consonant – a secondary articulation known as *pharyngealization*. This gesture underlies the "emphatic" consonants in Arabic (usually transcribed by underdotting or capitals), where we find such minimal pairs as *sayf* 'sword' vs. *ṣayf* 'summer' and *tiin* 'figs' vs. *ṭiin* 'mud'. According to Ghazeli (1977), the emphatics [ṣ,ṭ,ẓ,ð̣] are produced with a tongue body constriction in the upper pharynx.

The secondary articulations will be treated here as single segments rather than consonant-glide clusters. A number of considerations support this decision. First, many languages that otherwise lack consonant clusters freely permit C^w and C^y. Second, a secondarily articulated consonant may contrast with a consonant-glide cluster: compare Russian *s'est'* 'sit down' with *syest'* 'eat up'. Finally, when consonants with a secondary articulation simplify, the residue is typically a single segment. Two patterns of simplification will illustrate this. First, the primary articulation may be suppressed, leaving just the secondary articulation in the form of a glide. For example, in Polish the velarized [+back] lateral (orthographic *ł*) is now realized (in the standard dialect) as [w]: *Łódz* is [wuć]. Alternatively, the secondary place of articulation may be promoted to primary, displacing the original articulator. For example, the labialized velars k^w and g^w of Latin appear as bilabial stops in Romanian: Latin *aqua* 'water', *lingua* 'tongue' > Romanian *apă, limbă*. In this sound change, the [−continuant] feature is executed by the lips instead of by the tongue dorsum.

The most salient phonological property of the secondary articulations is their propensity to stretch over successive consonants or even syllables, mimicking the behavior of the corresponding features in vowel harmony. For example, in many Arabic dialects the pharyngealization of the emphatic phonemes spreads to other consonants of the root and beyond. In an analysis of the Lebanese dialect, Haddad (1983) shows that when the root [rxṣ] 'become cheap' maps to the *CiCiC* template, pharyngealization spreads from the final consonant to the remaining segments of the word, causing a nonfinal [i] to be realized as [u]: *ṛụx̣ịṣ* 'became cheap' vs. *rixis* 'became tender', built on the radical [rxs].

Consonants exist in which two articulators combine to form a single segment where the constriction of each is greater than a simple vowel or glide. The clearest examples of such *complex segments* are the *labiovelars* [kp] and [gb] found in many West African languages such as Ewe (Ladefoged 1975), with the minimal pairs shown in (42a).

(42)	a.	ekpã	'he faded'	egba	'he roofed'
		eka	'he chipped'	egã	'he became rich'
		epɔ	'he was wet'	eba	'he cheated'
	b.	epla	'he hurt'		
		eklɔ	'he washed'		
		ekpla	'he girded'		
	c.	fo	'to beat'	fofo	'beating'
		ci	'to grow'	cicii	'grown up'
		fle	'to buy'.	feflee	'bought'
		kplo	'to lead'	kpokplo	'leading'

These segments have the distribution of single consonants. The only clusters in Ewe consist of a consonant plus [l]; the labiovelar forms this cluster just like simple labial and velar stops (42b). Also, CV reduplication takes the first consonant and vowel of the stem (42c); once again, the labiovelar patterns as a single segment. When the labiovelars condition nasal assimilation, the result, at least in some languages, is a doubly articulated nasal [mŋ]. This suggests that the labial and velar components are unordered – at least phonologically. A controversial question in feature theory is whether one of the articulators must be singled out as primary. In the absence of such a distinction, we expect any rule mentioning labial or velar to be activated by the labiovelar complex. Not enough languages have been studied to know whether this expectation is justified.

Given that the labial and velar articulators may combine, it is natural to inquire about the behavior of the coronal. Sagey (1986) interprets the clicks of the Khoisan languages as dorsal consonants with coronal as a radically constricted secondary articulation. *Clicks* are formed by simultaneous constrictions of the tongue blade and dorsum, creating a temporary air chamber between the two closures. This chamber is rarified by sliding the tongue dorsum back along the soft palate, increasing the volume and consequently lowering the air pressure. When the coronal stricture is released, air flows into the mouth to create the distinctive click sound. These dramatic consonants are largely restricted to the languages of southern Africa, where they have been thoroughly integrated into the sound system. In Buru (Traill 1985) clicks combine with other manner-of-articulation features to produce some twenty distinct consonantal segments. While clicks are phonetically dramatic, their phonological behavior remains to be assessed.

Given that dorsal combines with both labial and coronal, we predict the existence of labiocoronals. Hoffman (1963) reports such consonants for Margi, citing *bdà* 'to sting (bee), to kick (donkey)', *bzə́r* 'child', *ptə́l* 'chief, king', *psár* 'grass'. He argues that these are single segments rather than clusters on the grounds that *ptə̀* 'to be insufficient' reduplicates as *ptə̀ptə̀* just as *sà* 'to drink' reduplicates as *sàsà*. As Maddieson (1987) points out, this argument is vitiated by examples such as *t͡sàgàlə̀*, which reduplicates as *t͡sàgàlət͡sàgàlə̀*, suggesting that the Margi reduplication is at the level of the entire stem and not just the first CV. If so, then the labiocoronals could be treated as consonant clusters instead of single segments. Sagey (1986) analyzes such clusters as *tk*ʷ, reported for certain Shona dialects by Doke (1931) as triply articulated single consonants. This interpretation has, however, been called into question.

1.11 Prosodic Features

The properties of length, stress, and tone are traditionally isolated from the other features into a special category of *prosodic* or *suprasegmental* features. Their distribution and phonological behavior characteristically ignores the features defining a sound's inherent quality. The autonomy of the suprasegmentals (and indeed the term itself) is reflected in the fact that orthographic systems register their presence (if at all) through diacritic marks or accents rather than with separate letters. The prosodic features have received intensive scrutiny during the past two decades. The insights acquired in the study of these categories have profoundly influenced phonologists' conception of the segmental features. We will review this research in considerable detail in later chapters. At this point we will content ourselves with a brief overview of the prosodic terrain.

Quantity

Many languages oppose long vowels and consonants to short. Phonetically, the duration of the longer sound exceeds that of the shorter by a factor of one-half to one-and-a-half or more. In consonants, length is usually indicated by *gemination* (doubling): Italian *nono* 'ninth' vs. *nonno* 'grandfather', Arabic *darasa* 'studied' vs. *darrasa* 'taught', Japanese *saka* 'slope' vs. *sakka* 'author'. Long vowels are variously transcribed as geminates (*taa*), by a colon (*ta:*), or by the macron (*tā*). Short vowels are unmarked (*ta*) or indicated by the breve (*tă*): Latin *mālus* 'apple tree' vs. *malus* 'evil', Arabic *kataba* 'wrote' vs. *kaataba* 'corresponded with', Japanese *tori* 'bird' vs. *toori* 'road'. Given the autonomy of suprasegmentals, quantity oppositions are usually distributed freely across a language's vowel or consonant inventory. There are, however, occasional gaps: for example, in Tiberian Hebrew geminates are excluded from the class of gutturals; in the Roman dialect of Italian palatals such as [č,š,ǰ,y] are phonologically geminate. In many systems length consists in the assignment of a phonological position after the syllable's nuclear vowel. This "slot" is then filled by either the following consonant or the preceding vowel. Such a conception of length affords a straightforward explanation for the behavior of the Hebrew definite prefix *ha*. When added to most stems, it geminates the following consonant: *seefer* 'book', *hasseefer* 'the book'; *melex* 'king', *hammelex* 'the king'. But when the noun begins with a guttural (which resists gemination), the vowel of the prefix lengthens instead: *ʔiiš* 'man', *haaʔiiš* 'the man', *ʕaam* 'people', *haaʕaam* 'the people'. Intuitively, the prefix has the phonological form *haX*, where *X* is an empty position that prefers to be occupied by the following consonant but when this is impossible