1

Natural classes and distinctive features in phonology

Most large cities in North America have an Ethiopian restaurant called "Blue Nile". A plausible explanation for this fact is that Blue Nile is an Ethiopian restaurant chain with franchises located throughout North America. There would be some evidence to back up this explanation: the restaurants have similar menus, some have similar decor and background music, and there even appear to be other Ethiopian restaurant chains with names such as "Queen of Sheba" and "Abyssinia" and locations across North America.

There is also some counterevidence to this claim. There is no independent evidence of a "Blue Nile Corporation": there is no web page, no publicly traded stock, no national advertising campaigns, and no pamphlet listing locations in other cities. Compared to known chains such as Starbucks®, it is apparent that different Blue Nile restaurants are implemented differently: they don't look the same, they don't use the same fonts on their menus, and their recipes aren't exactly the same. The food, decor, and music which are so similar between the restaurants are all things that are to be expected for an Ethiopian restaurant. Further, "Blue Nile" is a pretty good name for an Ethiopian restaurant, because the source of the Blue Nile is in Ethiopia, and many potential customers in North America have heard of the Nile River. It is not particularly surprising that multiple restauranteurs would choose the name "Blue Nile", so that approximately 7 percent of the Ethiopian restaurants in North America would share the name. In other words, the process by which entrepreneurs choose names for their Ethiopian restaurants favors names like "Blue Nile" over other logically possible names.

Other observations can be accounted for by looking at the factors related to the emergence of restaurant names: For example, the most frequent names are transparently "grounded" in potential patrons' awareness of Ethiopia. Second, even though Eritrean and Ethiopian food are quite similar, there are no Blue Nile Eritrean restaurants, for a very specific reason: the Blue Nile does not pass through Eritrea. The most common name shared by Ethiopian and

Eritrean restaurants is "Red Sea". Eritrea borders the Red Sea, and Ethiopia used to.[1] Concluding that there was an unseen entity Blue Nile Corporation would have prevented a deeper account of the similarities between the restaurants. The fact that "Blue Nile" is such a natural name for an Ethiopian restaurant suggests that it would be a good name for a chain, but the fact that it is such a good name is precisely the reason why it is not necessary to posit the existence of a corporation for which there is no direct evidence. The existence of many coffee shops named "Starbucks®" is better evidence for a Starbucks® Corporation because it is *not* a particularly natural name for a coffee shop, and therefore the corporation (for which there is quite a bit of direct evidence) is the only available explanation.

A parallel situation exists in language. Many languages in different parts of the world have similar sound patterns involving similar groups of sounds. A widely accepted explanation for this fact is that a small set of distinctive features which define these sounds are innate to humans. Theories of innate features have been used to account for many different observations about sound patterns. The purpose of this book is to argue that there are many sound patterns that innate features cannot account for, that there is no direct evidence for innate features, and that observations about sound patterns are better accounted for by emergent feature theory, a theory of how the development of sound patterns leads to the recurrence of particular groups of sounds, or natural classes.

1.1 Natural class behavior

Speech sounds in spoken languages do not always act independently. Instead, multiple sounds often participate in the same sound patterns. When a group of sounds exhibits the same behavior, it is often the case that these sounds are phonetically similar to each other. This type of grouping of sounds has been termed a "natural class", and the observation that phonological alternations often involve groups of sounds which share phonetic properties has led to the proposal that phonological alternations act upon specific properties of sounds, or "distinctive features", rather than on the sounds themselves. If a particular feature is targeted by an alternation, then all sounds bearing that feature are involved. Because many of the same groupings of sounds are observed in unrelated languages, it has been proposed that distinctive features are part of Universal Grammar, the innate and uniquely human capacity for

[1] Although there is no obvious synchronic motivation for naming an Ethiopian restaurant "Red Sea", there is a historical explanation: Ethiopia bordered the Red Sea before Eritrea gained independence in 1991, and Red Sea Ethiopian restaurants may all have been named at a time when there *was* synchronic motivation.

language. It follows from this that possible natural classes are those which can be characterized using the innate distinctive features. This has been a standard assumption in phonological theory since the 1960s.

For example, Turkish final devoicing applies not just to one type of sound, but to all of the non-nasal voiced consonants in the language, some of which are shown in (1). Consonants which are voiced word-medially are devoiced word-finally. Because devoicing is something that happens to all of these consonants in Turkish, it is claimed that the process applies not to segments, but to the feature [voice]. Final devoicing is observed in many unrelated languages, and this is taken as evidence that [voice] and other features are innate.

(1) Turkish final devoicing

a. Root-final nonnasal voiced consonants occur before vowel-initial suffixes.

kitab̬im	'my book'
kad̬im	'my floor'
fɛzim	'my fez'

b. These consonants are voiceless when word-final.

kitap	'book'
kat	'floor'
fes	'fez'

Distinctive features have been widely assumed to be part of Universal Grammar since the mid-twentieth century. While the theory of innate features predicts that a small set of distinctive features can describe most if not all natural classes, this prediction has never been explicitly tested. The usefulness of distinctive features in phonological analysis is clear from decades of research, but demonstrating that features are innate and universal rather than learned and language-specific requires a different kind of evidence. This book presents the results of the first large-scale crosslinguistic survey of natural classes. Based on data from 628 language varieties, the survey reveals that unnatural classes are widespread: among 6,077 unique classes of sounds which are targets or triggers of phonological processes in these languages, analyzed in three popular feature theories (*Preliminaries to Speech Analysis*, Jakobson et al., 1952; *The Sound Pattern of English* (*SPE*), Chomsky and Halle 1968; and Unified Feature Theory, Clements and Hume 1995), no single theory is able to characterize more than 71 percent of the classes, and over 24 percent are not characterizable in *any* of the theories. While other theories are able to account for specific subsets of these classes, none is able to

predict the wide range of classes which actually occur and recur in the world's languages.

This book argues that the natural classes and distinctive features found in human languages can be accounted for as the result of factors such as phonetically based sound change and generalization, which can be described without reference to a feature system. A feature system can be constructed (by a language learner or a linguist) on the basis of the results, but the feature system critically does not need to be a driving force behind sound patterns. Facts which have been attributed to innate features are accounted for by independently needed concepts (such as language change and similarity). It follows that phonological distinctive features no longer need to be assumed to be innate.

It is no secret that there are phonological patterns which do not conform to models of innate features, and a common approach is to treat these as marginal processes which are beyond the purview of innate feature models. One example is palatalization in the Chi-Mwi:ni dialect of Swahili (Kisseberth and Abasheikh 1975, Clements 1985), in which certain consonants undergo palatalization before the perfect suffix -i̱ɬ-. The only place feature these consonants retain their value for is *SPE*-era [anterior]. [g] is an exception, because it loses its value to change to [z], instead of the expected [ʒ] (2).

(2) p t̪ t → s
k → ʃ / [+nasal]__
bd̪dgɬ → z

This is problematic for innatist approaches which hold that all place features are expected to spread as a constituent. Rules such as the one in (2) appear to be the result of telescoping (the merging of independent rules), and Clements (1985: 246) draws a distinction between this type of rule and those which are captured simply using innate features and feature organization:

> We will not relax the empirical claims of our theory in order to provide simple descriptions of rules such as these, since if we did so we would fail to draw a correct distinction between the common, widely recurrent process types that we take as providing the primary data for our theory, and the sort of idiosyncratic phenomena whose explanation is best left to the domain of historical linguistics.

The strongest versions of innate feature theory might require the relationship between attested or attestable phonologically active classes and featurally natural classes to be identity. However, many phonologists may not expect features to predict all of the classes that occur, because many of the classes are the historical residues from millennia of diachronic changes that may interfere

with naturalness. The telescoping of multiple natural changes may result in seemingly arbitrary unnatural-looking synchronic patterns. While these idiosyncratic classes might be treated as fundamentally different from natural classes, there is also a relationship between natural changes and natural classes. Recurrent natural classes may also be understood as simply the most common types of historical residue.

In accounting for unnatural classes, it makes sense to separate the classes into two categories: those which are phonetically natural and may have a transparent phonetic basis, but have the misfortune of being natural according to a set of properties for which a distinctive feature has not been proposed, and those which are phonetically unnatural, and likely arose via a series of changes, each of which may have been natural but whose end result is a phonetically unnatural class. One approach is to aggressively expand the innate feature set to account for all phonetically natural classes while forsaking unnatural classes. Anderson (1981) argues that this can be a very difficult distinction to make, because many classes which appear to have a transparent basis in phonetically natural changes turn out to have very different origins. Further, a vast theory of innate features which attempts to capture all phonetically natural classes without adopting an independent gradient notion of naturalness would need to draw a distinction between the most marginally natural (but perhaps unattested) classes which are admitted and classes for which there is no apparent phonetic basis (but which may be attested nonetheless) which are rejected.

Another approach, taking into account the experimental evidence that natural and unnatural classes may be processed the same way synchronically, is to treat all classes as historical residues, and to explore phonologically active classes in terms of their historical development, to understand the types of class a learner may be called upon most frequently to acquire. In this view, natural classes are a special case of idiosyncratic historical residues, i.e. they are the ones which most transparently reflect their phonetic origins and which therefore occur most frequently and are most likely to be encountered and embraced by phonologists.

While historical explanations are often invoked within innate feature approaches in order to account for problematic cases, it is unclear how often such an explanation can be invoked. Is it a coincidence that a model of synchronic phonology such as innate features is well suited to modeling processes which commonly arise from phonetic motivations (and for which a straightforward phonetic explanation exists) and ill-equipped to model less common phonological processes (for which only a more complicated phonetic explanation exists)?

Suppose that explanation from innate distinctive features is a medium-sized rectangle, and that explanation from phonetics and language change is a large triangle. The argument that phonological processes can be explained by innate distinctive features (phonetically grounded or not) amounts to the statement represented in Fig. 1.1.

Suppose that a sample of phonological processes includes examples (such as the one from Chi-Mwi:ni Swahili) that fall outside the rectangle (Fig. 1.2). The rectangle has already been placed very carefully by many bright linguists, on the basis of observed patterns and phonetic considerations, but counter-examples persist. Counterexamples that appear to defy innate features are often argued to be beyond the purview of the feature system, and have been accounted for by invoking external factors such as language change and physiology. Accounting for these by invoking external factors amounts to adding extensions (small triangles) to account for problem cases (Fig. 1.3).

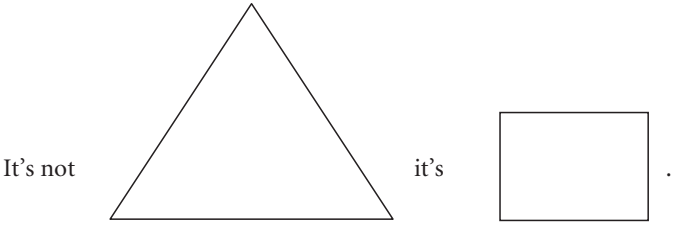


FIGURE 1.1 Factors vs. features

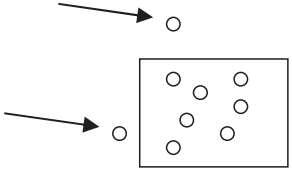


FIGURE 1.2 Innate feature theory with exceptions

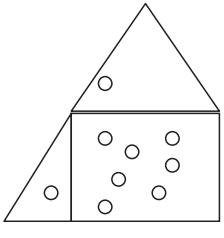


FIGURE 1.3 Innate feature theory with extensions

Adding small triangles to the rectangle causes it to look more and more like the triangle that was rejected in the first place, and the distinction between the modified rectangle and the larger triangle is less useful. It is argued in this book that sound patterns can be accounted for more effectively by dispensing with the rectangle/triangle distinction, i.e. by treating recurrent, natural classes as a special case of historical residues (the residues favored by the factors relevant to the development of residues) rather than a separate privileged category.

1.2 Emergent feature theory

In showing that innate distinctive features are unnecessary to explain the existence of natural classes, it is not necessary to deny that features are a relevant part of a phonological system. Features which arise in the way proposed here are just as well suited as innate ones for defining phonological patterns, forming contrasts, and doing everything else that features have been claimed to do. Emergent feature theory simply offers a different explanation for the existence of phonological features, one which is more compatible with knowledge of genetic and linguistic change, and with known synchronic phonological patterns.

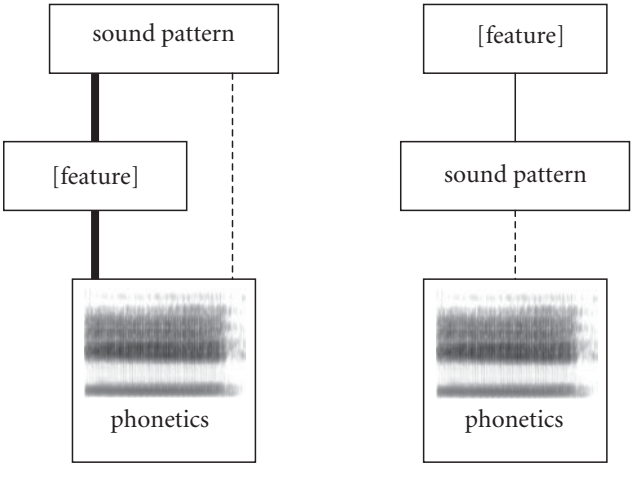
Emergent feature theory is at least partially consistent with and/or inspired by a good deal of work previous work on:

- the emergence of linguistic patterns and structure (e.g. Martinet 1968, Lindblom 1983, 1984, 1986, 1990a,b, 1999, 2000, Ladefoged 1984, Bybee 1985, 1998, 2001, Corina and Sagey 1989, Port 1996, Steels 1997, MacWhinney 1998, Pierrehumbert 2001, 2003, Beckman and Pierrehumbert 2003, Pulleyblank 2003, and Wedel 2004);
- the unnaturalness of some synchronic patterns (e.g. Bach and Harms 1972, Anderson 1981, Ladefoged and Everett 1996, Buckley 2000, and Hale and Reiss 2000);
- the explanation of synchronic observations in terms of diachrony (e.g. Andersen 1972, 1973, Anttila 1977, Ohala 1981, 1983, 1992, 1993a,b, 2003, Labov 1994, 2001, Newmeyer 1998, Blevins and Garrett 1998, Garrett and Blevins 2004, Dolbey and Hansson 1999, Janda 1999, 2001, 2003, Hyman 2001, Janda and Joseph 2001, 2003, Myers 2002, Vaux 2002, Guy 2003, Hale 2003, Kiparsky 2003, Blevins 2004, and Culicover and Nowak 2004);
- the explanation of synchronic observations in terms of external factors (Beddor 1991, de Boer 2000, Hume and Johnson 2001a, Kochetov 2002, and Hume 2004a, b); and

- approaches to morphology as a distinct component of grammar (e.g. Maiden 1992, Aronoff 1994, Carstairs-McCarthy 1994).

There is a general trend in the field toward narrowing the scope of the uniquely human language faculty, typified by Hauser et al. (2002). The idea that phonological classes are language-specific is consistent with language development-based arguments that phonological (Vihman and Croft 2007) and grammatical classes (Croft 2001, Tomasello 2003) are emergent.

Innatist and emergentist approaches both posit relationships between phonetic substance, abstract features, and the phonological patterns found in human languages. The difference lies in the nature of these relationships. For innate features (Figure 1.4a), abstract features are grounded directly in phonetics, and phonological patterns reflect both the features and the phonetic substance because features are the building blocks of phonological patterns. The relationship between phonological patterns and phonetics (bypassing features) is less direct, but necessary in order to provide the phonetic or historical accounts for "idiosyncratic" phenomena which are difficult or impossible to analyze with the given features. For emergent features (Figure 1.4b), this loose relationship between phonetics and phonological patterns is the sole connection between phonological patterns and phonetic tendencies. Just as grammar-external factors (including external phonetic factors) can be used to account for idiosyncratic phenomena in an



a. innate features b. emergent features

Figure 1.4 Relationships between phonetics, features, and phonological patterns

approach which otherwise depends on innate features, phonetics can account for these unusual phonological patterns, and *also* for more common patterns, which tend to reflect more common phonetic tendencies. In this way, emergent feature theory employs a single mechanism to account for common and rare phonological patterns, in contrast with innate feature theory, which employs two.

In emergent feature theory, features are abstract categories based on generalizations that emerge from phonological patterns. In innate feature theory it is typically the other way around: phonological patterns emerge (at least in part) from the effects of features. Whereas innate features are typically grounded directly in phonetics, this relationship is different for emergent features: recurrent phonetically defined features reflect phonetics *via* the phonetically grounded phonological patterns they are motivated by. Because features are abstract, there need not always be a connection between phonetics and phonological patterns, and features do not necessarily always refer to phonetically natural classes.

A more detailed view of the relationship between features, phonological patterns, and external factors is given in Chapter 5. The environment in which language is used includes the anatomy used to produce and perceive speech, the laws of physics this anatomy is governed by, the social context in which language is used, and the cognitive mechanisms employed in learning and using language. The factors audition, attention, categorization, aerodynamics, coordination, and social identity contribute to the development of the phonological patterns found in language, making some patterns more common than others. The influence of these factors on sound patterns is illustrated in Fig. 1.5. The features which learners use to define these sound patterns reflect the factors that influence them, and have the potential to influence the patterns (the reason for the bidirectional arrow). The role of speech production and perception is not to be interpreted as simply ease of articulation and ease of perception, but as the physiological and cognitive realities in which language exists. The external factors in Fig. 1.5 and their relationships with each other, sound patterns, and features will be discussed in more detail in Chapters 5 and 8.

Emergent feature theory holds that phonetic factors shape the phonological patterns of the world's languages, and these patterns can be internalized by speakers in terms of features which are necessary to describe them, rather than in terms of predetermined innate features. These external influences lead to classes which tend to involve phonetically similar segments. The use of phonetically defined distinctive features is just one way to describe classes of phonetically similar segments. While these types of explanation are often invoked to account

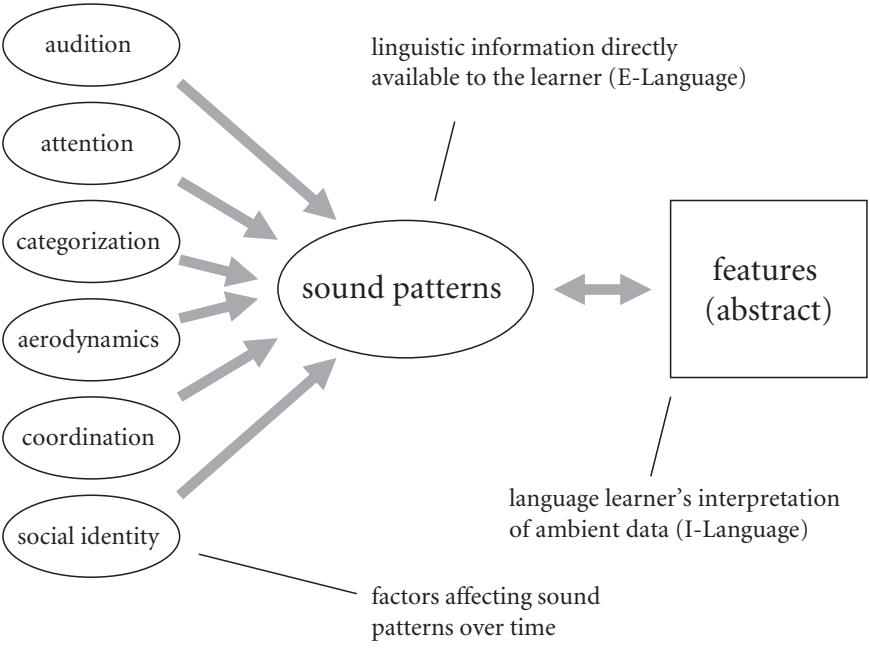


FIGURE 1.5 Abstract features from concrete external factors

for "idiosyncratic" unnatural classes, it will be shown that they are even better at accounting for "natural" classes, and the result is a unified account of what have previously been considered to be natural and unnatural classes.

1.3 Incorporating insights of innate features into emergent feature theory

Innate feature theory has captured many different insights about phonological patterns over the years. Because it abandons the assumption of innateness, emergent feature theory needs to account for these observations in other ways. One important observation is that there are a limited number of phonetic parameters available for use in language. Jakobson et al. (1952) suggest that all languages can be described using the twelve oppositions vocalic/non-vocalic, consonantal/non-consonantal, interrupted/ continuant, checked/unchecked, strident/mellow, voiced/unvoiced, compact/diffuse, grave/acute, flat/plain, sharp/plain, tense/lax, and nasal/oral. It will be seen below how well these features are represented in the phonologically active classes of 628 language varieties. To the extent that there is a crosslinguistic

preference for these oppositions, emergent feature theory accounts for them in roughly the same way Jakobson et al. account for them, by observing that there are a limited number of phonetic parameters available to language, and that phonological patterns reflect that. Jakobson et al.'s features are stated in acoustic terms, but they observe also that the acoustic parameters associated with these features correspond to specific articulatory parameters, and they account for typological observations in terms of these parameters. For example, they account for the apparent absence of languages which contrast pharyngealization and labialization separately by noting the acoustic similarity of the two types of articulatory gesture, and consequently allow the feature [flat] to represent the acoustic property that is produced by two different articulatory means. It is now known that there are languages such as Tamazight Berber (Abdel-Massih 1968) with contrastive pharyngealization and labialization, but the finding that the coexistence of these contrasts is much rarer than the coexistence of many other contrasts still stands.

Emergent feature theory attributes the *rarity* of such languages to acoustic *similarity*, and attributes the *possibility* of coexistence to the articulatory difference and acoustic *non-identity*. Because it uses similarity to predict the likelihood of phonological patterns, emergent feature theory is better equipped to distinguish between similarity and identity than is innate feature theory. In formulating linguistic theories, it is very tempting to identify similarity with identity. The upside of confusing similarity with identity is that it allows more sweeping generalizations to be made. The downside is that they are often wrong.

Another observation is that articulatory parameters are relevant to phonology. It has been proposed (e.g. Chomsky and Halle 1968, Sagey 1986) that all phonological patterns can be accounted for with an innate set of articulatory features. In *The Sound Pattern of English* (*SPE*), the features themselves, rather than phonetic parameters, are the explanation for observed phonological patterns. Emergent feature theory accounts for the same observations on the basis of language change, phonetic similarity, and the cognitive process of generalization. As shown in Chapter 7, the classification of phonologically active classes involves many of the articulatory parameters identified by Chomsky and Halle, as well as parameters they do not identify.

A third observation to be accounted for is that some phonetic parameters are interdependent on each other, and some act independently. This is represented in Feature Geometry (e.g. Clements 1985, Sagey 1986) and Dependency Phonology (Anderson and Ewen 1987, Harris 1994) by a feature hierarchy with constituents which correspond to features that pattern together. Features which are linked under the same node tend to be features

which are linked articulatorily. In this way, Feature Geometry is an abstract model of some of the phonetic parameters relevant to phonology. In abstracting away from the phonetic basis for phonology, the different versions of Feature Geometry highlight some of the phonetic parameters which are most important for determining phonological patterns as well as the ways in which they interact with each other.

1.4 Definitions

The term "natural class" is used to mean different things, and it will be necessary to be precise about how the term is used in this book. The traditional definition has two parts, as in (3).

(3) Natural class (traditional two-part definition)
   i. A group of sounds in an inventory which share one or more distinctive features, to the exclusion of all other sounds in the inventory.
   ii. A group of sounds in an inventory which may participate in an alternation or static distributional restriction, to the exclusion of all other sounds in the inventory.

These two definitions are often assumed to be equivalent, and if it can be demonstrated that phonological alternations do indeed act only upon distinctive features, then these definitions would be equivalent. Because one of the goals of this study is to find out if the two definitions really are equivalent, this is not something that will be assumed. When the term "natural class" is used in the rest of this book, it will be used in terms of a particular feature theory, using the theory-dependent definition in (4).

(4) Natural class (feature theory-dependent definition)
A group of sounds in an inventory which share one or more distinctive features *within a particular feature theory*, to the exclusion of all other sounds in the inventory.

It is often assumed that that phonological natural classes are phonetically natural, as defined in (5). In cases where this is the intended interpretation, the term "phonetically natural class" will be used instead.

(5) Phonetically natural class
A group of sounds in an inventory which share one or more phonetic properties, to the exclusion of all other sounds in the inventory.

Note that this definition is broader than the one in (4), because not all phonetic properties have features assigned to them in each theory. An

"unnatural class" is a class that does not meet a particular set of criteria for being natural. What has been dispensed with in the definitions in (4, 5) is any reference to phonological patterning, which is crucially not *assumed* to be identified with phonetic similarity or shared features. To refer to classes which participate in phonological patterns, the term "phonologically active class" will be used. This term is defined in (6). It is a crucial point that while any phonologically active class is, by definition, naturally occurring, there is no guarantee that it is a "natural class" with respect to any given feature theory (4) or "phonetically natural" with respect to any interpretation of phonetic similarity (5).

(6) Phonologically active class (feature theory-independent definition)
A group of sounds in an inventory which do at least one of the following, to the exclusion of all other sounds in the inventory:
- undergo a phonological process,
- trigger a phonological process, or
- exemplify a static distributional restriction.

It is useful to give some examples of these different types of class. For example, Japanese has a well-known sound pattern in which unaccented high vowels are devoiced between voiceless consonants and word-finally after the same consonants, as in (7).

(7) Japanese vowel devoicing (Vance 1987, Shibatani 1990)[2]

/kɯtsɯ'/	→	[kɯ̥tsɯ]	'shoes'
/ha'ʃi/	→	[haʃi̥]	'chopsticks'
/sɯsɯki'/	→	[sɯ̥sɯ̥ki]	'eulalia'

Because the the vowels /i ɯ/ are targeted by this sound pattern and the consonants /p t k s ʃ h/ trigger it, both of these are phonologically active classes, and provide data points for the survey. Additionally, both of these are featurally natural classes with respect to the feature system of *The Sound Pattern of English* (Chomsky and Halle 1968), among others, because they can be represented with the conjunction of one or more features, in this case [+high, +tense] and [−voice]. Finally, they are both phonetically natural classes as well, because they can be described, to the exclusion of all other sounds in the inventory, in terms of measurable phonetic parameters such as second formant frequency, duration, and vocal fold vibration.

The consonants /t k s ʃ h/ are another phonologically active class in Japanese, because they are voiced when they appear initially in the second

[2] Note that the final vowels in 'shoes' and 'eulalia' are not devoiced because they are accented.

part of compounds, a pattern known as "sequential voicing" or *rendaku* (8a). /p/ is alone among the voiceless consonants in not participating in sequential voicing (8b).

(8) Japanese sequential voicing (Vance 1987, Shibatani 1990: 173–4, McCawley 1968: 187)

a.	[ama]	+	[tera]	→	[amadera]
	'nun'		'temple'		'nunnery'
	[ɯwa]	+	[tsɯmi]	→	[ɯwadzɯmi]
	'over'		'piling'		'upper load'
	[oo]	+	[same]	→	[oozame]
	'big'		'shark'		'big shark'
	[to]	+	[ʃimaɾi]	→	[todʒimaɾi]
	'door'		'closing'		'locking of a house'
	[tabi]	+	[hito]	→	[tabibito]
	'travel'		'person'		'traveler'
	[iɾoha]	+	[kaɾɯta]	→	[iɾohagaɾɯta]
	'Japanese syllable counting'		'cards'		'playing cards with *hiragana* on them'
b.	[genmai]	+	[pan]	→	[genmaipan]
	'whole rice'		'bread'		'whole rice bread'

While this is a phonologically active class, it is not a featurally natural class in *SPE*, because there is no conjunction of *SPE* features which can describe the set of all voiceless consonants except /p/. It is a phonetically natural class, though, because it can be described in terms of measurable properties such as lip closure, constriction degree, and vocal fold vibration. An analysis of this sound pattern in *SPE* features requires devices beyond the scope of the feature system, such as rule ordering or antagonistic constraints or conjunction of feature bundles (i.e. bracket notation).

Finally, the consonants /b d z n m ɾ/ are a featurally natural class in Japanese with respect to *SPE*, because they can be described with a conjunction of features such as [+voice, +anterior], and a phonetically natural class, because they can be described in terms of measurable properties such as stricture location and vocal fold vibration. However, this is not a phonologically active class in Japanese, because there is no reported sound pattern which targets or is triggered by specifically these segments.

With these definitions in hand, it is now possible to proceed to investigating the connections between these different types of class, and how they might be accounted for. The concepts "phonologically active class", "featurally natural

class", and "phonetically natural class" have often been conflated in phonological theory as the monolithic "natural class". There is obviously a considerable amount of overlap because featurally natural classes tend to be phonetically natural (because most if not all widely accepted features are phonetically defined) and because many phonologically active classes are natural both featurally and phonetically. It is clear that there are many phonetically and featurally natural classes which are not phonologically active in a given language or perhaps any language, and the survey results will show that many possible phonologically active classes are also not featurally or phonetically natural, and others are phonetically natural without being featurally natural. Due to the considerable overlap between phonetic and featural naturalness, it is difficult to assess how much responsibility each bears for the nature of phonologically active classes.

1.5 General arguments against innate features

Beyond phonological evidence which is the subject of Chapters 4, 6, and 7, there are many reasons to be suspicious of the idea that distinctive features are innate. In this section, arguments from signed languages and from phonological theory are presented, pointing to the conclusion that features are not universal or innate. The purpose of this discussion is not to underestimate the contribution of innate feature proposals to our understanding of phonological systems, but to examine the specific proposal that distinctive features are innate. While innate features are central to the way most of these approaches to phonology are implemented, the insights about phonological patterning which have been cast in terms of innate features in the past fifty years stand on their own, and emergent feature theory could not proceed without them.

1.5.1 *Signed language features*

Most work in feature theory focuses on spoken languages, and typological surveys, markedness generalizations, and hypothetical universals are generally made on the basis of only spoken language data. While substantial work has been conducted in the area of sign language features and feature organization (e.g. Stokoe 1960, Liddell 1984, Liddell and Johnson 1989, Sandler 1989, Brentari 1990, 1995, 1998, Perlmutter 1992, van der Hulst 1995, Uyechi 1996), there are obvious practical reasons for focusing on a single modality (and the survey in this book only includes spoken language data). Focusing on spoken language allows modality-specific questions to be addressed (such as the role

of the vocal tract and auditory system in phonology), but questions about phonological universals cannot ignore the existence of sign language phonology.[3]

The hypothesis that there is a small set of innate distinctive features which are defined in terms of the articulation and/or audition of spoken language and which are the only features available to the phonologies of the world's languages is incompatible with signed language phonology, because signed languages involve an entirely different set of articulators and rely primarily on vision rather than on audition. Consequently, the claims about an innate feature set must be qualified with the acknowledgement that this universality is really only applicable to languages of one modality, even though UG purportedly applies to all languages.

There are a number of ways to reconcile the universalist claims with the existence of signed language phonology: (1) relax the requirement that features are defined in phonetic terms and interpret each innate feature as having both spoken language and signed language phonetic correlates, (2) posit additional innate features which apply to signed language, and claim that humans are hardwired with two sets of innate features for two different modalities, or (3) consider that features and their phonetic correlates are learned during acquisition, according to the modality of the language being acquired.

If signed and spoken languages use the same innate features but with different phonetic correlates, it is expected that there would be some evidence that they are otherwise the same features. This evidence could include feature geometries for signed languages that look like Feature Geometries for spoken language. Research in signed language features offers no such evidence (see Brentari 1995, 1998 for reviews). In fact, Liddell (1984) reports that evidence from American Sign Language suggests that signed languages have significantly larger numbers of contrastive segments than spoken languages, and many other analyses are consistent with this. Stokoe (e.g. 1960) produced the first phonemic analysis of signed language, using 12 distinctive places of articulation, 18 distinctive handshapes, and 24 distinctive aspects of movement. The Hold-Movement Model (e.g. Liddell and Johnson 1989) involves 299 distinctive features. Brentari (1990) reorganizes Liddell and Johnson's feature system and reduces the number of features to 20, a number more comparable to that proposed for spoken languages; but Brentari's analysis achieves this by using seven features with more than two values, in addition to other binary and privative features.

[3] The importance of considering sign languages when formulating linguistic universals is discussed further in Blevins (2004: 301–4) and Sandler and Lillo-Martin (2006, esp. 272–8). Haspelmath et al.'s (2005) *World Atlas of Language Structures* includes a survey of morphosyntactic properties of signed languages (Zeshan 2005).

Sandler's (e.g. 1989) Hand Tier model was the first to incorporate a hierarchical organization of features, placing hand configuration and location on separate trees, as shown in Figs. 1.6 and 1.7, and bears little resemblance to any spoken language Feature Geometry proposals. Similarly, other feature organizations such as the Dependency Phonology model (e.g. van der Hulst 1995), Visual Phonology (e.g. Uyechi 1996), or the Moraic Model (e.g. Perlmutter 1992) do not resemble proposed spoken language Feature Geometries.

The similarities between the feature organizations for different modalities are limited to very general statements, such as the observation that both have a place node. Just as spoken language feature organization reflects the physiology of the vocal tract, signed language feature organization (e.g. as seen in Figs. 1.6 and 1.7) tends to reflect the anatomy that is relevant for signed language. For example, the organization of features in the Hand Configuration tree, such as the features [T], [I], [M], [R], [P], representing fingers, corresponds to the organization of body parts. Beyond the representation of physiology in feature hierarchies (as is seen in spoken language), Brentari (1998) draws parallels between the structure of signed language phonology and the human visual system, just as many sound patterns in spoken languages reflect the human auditory system. If features are driving phonology, and these are the same features, there should be observations that are attributable only to the features and their organization, rather than to commonalities between the physiological facts they represent.

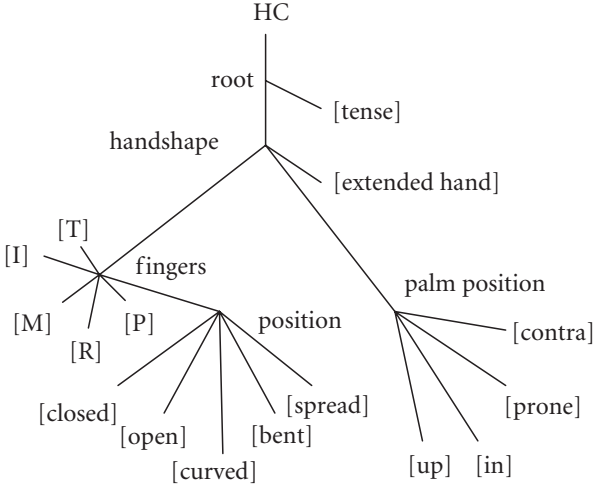


Figure 1.6 The Hand Configuration tree (Sandler 1989)

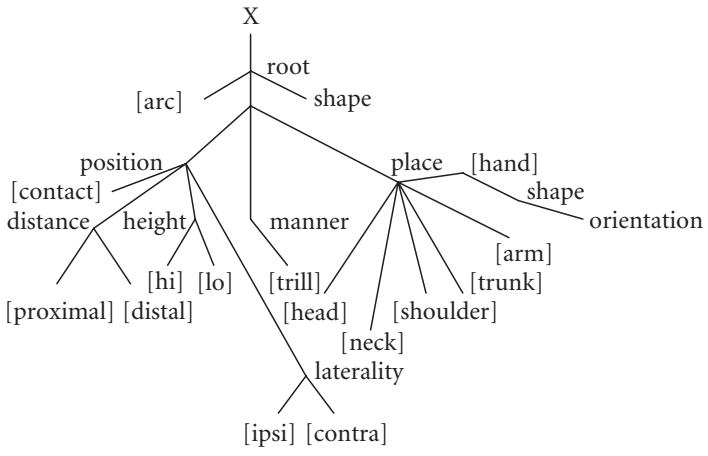


Figure 1.7 The Location tree (Sandler 1989)

Positing separate feature sets for signed and spoken languages runs into specific problems, namely that even if spoken language features could have evolved through natural selection, it is not very plausible that signed language features did as well, because most humans are not deaf, and because deafness is rarely hereditary. It is not clear how a genetic mutation introducing an innate signed language distinctive feature could have been advantageous before Deaf communities became established in fairly recent times (e.g. the first Deaf school was established in the 1500s).

If phonetic correlates, and perhaps feature organization, are assigned by the language learner in acquisition, then what is shared by signed language and spoken language phonology may simply be cognitive categories. In other words, categories/features emerge as a result of contact with language data, and they naturally reflect the modality of the language being learned. A child learning a signed language will develop features associated with the production and perception of signs, and a child learning a spoken language will develop features associated with the production and perception of speech. This is essentially the position taken by Brentari (1998: 313) regarding differences between signed and spoken language: that the formal role of distinctive features and other primitives is the same for both modalities, but that the substantive definitions of them depend on modality and experience.

The idea that signed language features, and thus perhaps all features, must be learned, is not new. Corina and Sagey (1989) analyze phonological alternations in ASL using a feature-geometric framework. They note that the

proposed Feature Geometry for signed languages is clearly related to anatomy and very different from the Feature Geometry models proposed for spoken languages (e.g. Sagey 1986), which are also clearly related to anatomy (but different parts). Finding it implausible that signed language features are in UG, they try to reconcile the differences between the two:

> An alternative [to putting them in UG] is to say that sign language hierarchies are learned or derivable from some language external facts. Since the features and the feature hierarchy are closely tied to articulation, this is not an implausible result. In fact, their being learned could explain why they are clearly tied to articulation. But we are left with a peculiar state of affairs. We posit an innate feature system for spoken language, but a derivable one for signed languages. Once again this seems inconsistent. Could it be the case that spoken language features and hierarchies too are derivable or learned constructs rather than innate? If we adopt this position that features and feature hierarchies are learnable and not given in UG, we open up the possibility that they are not completely universal. That is there could be slight differences between languages, the particular language influencing the feature set and the hierarchy. The vast differences in the feature hierarchy proposed here simply represent the extreme end of this continuum, due to the radically different mediums in which they are conveyed. The puzzle to be explained would now become why hierarchies are so similar among languages. If features are in UG, then any variations must be explained; if features are not in UG, then any universals among languages must be explained. (Corina and Sagey 1989: 81–2)

During the past half-century, phonologists have generally taken UG as the explanation for crosslinguistic similarities, and sought special explanations for apparent exceptions. Emergent feature theory takes the opposite approach. The fact that features and feature hierarchies appear to be so similar may not be so much a puzzle as a result of the *assumption* that features and feature hierarchies are so similar. In fact, as will be seen in later chapters, most languages do not have phonological phenomena to motivate most features. There is no reason to believe that these languages have particular features except for the assumption that all languages must have the features which are motivated by other languages. The differences between the features which are useful for analyzing signed and spoken languages demonstrate how much the similarities are dependent upon modality.[4]

[4] It is worth noting that the survey data presented in subsequent chapters of this book is from spoken languages. In the discussion, the term "sound pattern" is generally used in contexts which for some reason do not apply to signed languages. The term "phonological pattern" is used in more general contexts where exclusion signed languages is not intended.

1.5.2 *No evidence that unattested = impossible*

The goal of many theories of phonology is to distinguish possible phonological phenomena from impossible ones. Often the only evidence given for the impossibility of a phonological pattern is that it is unattested in the fraction of existing spoken languages which have been described; for example, Sagey (1986: 9) writes: "It should be possible to represent within the theory any phonological process or form that is possible in human language, and it should be impossible to represent phonological forms and processes that do not exist in human language."

The ability to represent all and *only* the phonologically active classes which recur is described by McCarthy (1994: 191) as the most import criterion for an adequate theory of distinctive features (emphasis mine):

> An adequate theory of phonological distinctive features must meet four criteria: (i) it must have a relatively consistent and direct relation to the phonetic properties of speech sounds; (ii) it must be able to describe *all and only the distinctions* made by the sound systems of any of the world's languages; (iii) it must be able to characterize *all and only the natural classes of sounds* that recur in the phonological phenomena of different languages; and (iv) it must correctly characterize the subgroupings of features by recurrent phonological phenomena. The third criterion is the most important one and probably the hardest to achieve.

At least two questions are relevant here: First, how confident are we that phonological patterns which are unattested in today's documented languages are impossible? The number of languages which have been documented are a small sample of the languages which exist, and the number of languages which are currently living are just a small sample of the languages which have existed and will exist in the future. When there are so many linguistic phenomena found in only a handful of attested languages, how can we be certain that *any* phonological pattern never existed in the past, never will exist in the future, and doesn't exist currently in an understudied language?

Chomsky and Halle (1968: 4) contrast linguistic universals and accidental universals. To illustrate accidental universals, they construct a hypothetical scenario in which only inhabitants of Tasmania survive a future war. In this scenario, it would be a true generalization to say that no existing language uses pitch to distinguish lexical items; but Chomsky and Halle argue that this would be useless information to linguistic theory, because this generalization is only true by virtue of the elimination of most of the world's population by a non-linguistic event.

War, genocide, and other events have already destroyed entire language families. Phonological patterns that were unique to these languages are

unattested in today's languages, and making it impossible to represent them inevitably rules out possible forms that the human language faculty is capable of dealing with (and has dealt with before). Theories of representations which exclude unattested patterns are valued in many approaches to feature theory and phonetically driven phonology, and this is a common assumption in Optimality Theory (factorial typology).

Whether or not the phonological formalism should rule out unattested phonological patterns is a very important issue. While it is clearly important to have a theory of possible and impossible or likely and unlikely phonological phenomena, there is no reason to believe that the formalism for the cognitive representation of phonological patterns is the only appropriate place for such a theory.

One of the reasons for positing a *small* set of innate features is to keep the theory from overgenerating, i.e. being able to represent phonological patterns which have not been observed. The languages which have been documented give a picture of what types of phonological pattern are expected; it is justified to conclude that phonological patterns which occur frequently in the sample are common crosslinguistically. However, if a pattern is unattested in documented languages, it is not justified to conclude that it is impossible. This is because there are so many phenomena which are attested only once, and which the same criteria would likely deem impossible if a different sample were selected. While it may be justified to conclude on the basis of a sample that a pattern is rare, there is a major difference between rare and impossible when the issue is whether the language faculty should be capable of dealing with the pattern at all.

1.5.3 *No null hypothesis and no large-scale survey*

While most feature theories are supposed to cover all spoken languages, the arguments in favor of particular versions of innate feature theory generally consist of examples from a handful of languages which are dealt with in an elegant fashion by the theory being advocated. The success of a given feature theory, combined with the assumption that features are innate, is taken to support the assumption that features are innate and to validate the model in question. The fact that a variety of feature theories are able to account for different phonological phenomena using phonetically defined features is consistent with the idea that a variety of phonetic facts are relevant for accounting for phonological phenomena. Even if they conflict, it is not surprising that there are many different competing theories of innate features, since each one is valid for some set of data but lacks the ability to account for data that some other theory is better suited for. The claim that one theory in particular is innate and universal is a leap that requires the evidence that

would be provided by a large-scale survey. In addition to competing feature theories, it is quite common for feature systems to be tweaked slightly in specific cases in order to better fit the data. Not considering all of these tweaks at once gives the false impression that conventional feature systems are handling all of the data. Considering a wide array of languages all at once suggests that phonological theory should recognize a feature system which is officially as malleable as feature theory has been in practice all along.

McCarthy (1994: 191, quoted above) describes the ability to characterize all recurrent natural classes as the most important criterion for an adequate theory of distinctive features (and the most difficult to achieve), and it will be seen in the survey results that all universal feature theories fail to meet this criterion.

Aside from the optimistic goal of accounting for everything, there is no theory of *how much* phonological patterning should be accounted for by a feature theory in order to motivate the innateness of its features. Arguments for innate feature models do not involve a theory of the extent to which phonetic factors would be expected to influence phonology *anyway*, without the existence of an innate feature set. A possibility that is generally ignored is that the successes of a given model of features can be taken as evidence that the model is correct in its choice of articulatory and acoustic facts to recapitulate, but in itself unnecessary precisely because these phonetic explanations already exist.

Innate feature are often treated as though they are the primary explanation for the fact that sound systems in different languages tend to resemble each other, as though they would be in chaos without being regimented by innate features. For example, Clements and Hume (1995: 245) state that feature theory "explains the fact that all languages draw on a similar, small set of speech properties in constructing their phonological systems" and "has provided strong confirmation for the view that languages do not vary without limit". In the view that features explain why sound systems do not vary without limit, the similar patterning of speech sounds is taken as evidence that there are universal features. It is probably uncontroversial, though, that sound patterns would not really be in complete chaos without innate features, that the null hypothesis is not that all logically possible phonological patterns should be equally likely in human language. For example, "/car horn/ → [60 Hz hum] / __fruit bat chirp" is a logically possible phonological pattern. Even without innate features, the absence of this pattern can easily be attributed to the fact that the human vocal tract is not well suited to producing these sounds, the human auditory system cannot detect them all, and that even if the sounds were producible and perceivable, it is unclear what

diachronic changes would lead to such an alternation. The null hypothesis must take into account the fact that the speech sounds of human spoken languages are limited by human physiology and general cognitive capacity, and that natural languages are not invented by their speakers but descended along sometimes familiar paths from earlier languages. Given this, the case for an innate feature set could be strengthened by specifying the minimum amount of similar patterning that must be found, and what its nature must be, in order to conclude that an evolutionary leap has created an innate feature set. The same applies to extragrammatical features of language use which are presented as arguments for an innate feature set. What would we expect language acquisition, disablement, and change to look like in a world without innate features but with familiar tangible constraints on possible languages?

In addition to the lack of a null hypothesis with which to compare innate feature theories, there have been no large-scale typological studies examining the predictions of various models. This book provides the results of a large-scale typological survey in order to examine the extent to which innate feature theories and the phonetic factors they are grounded in are able to account for phonological patterning in a wide range of languages.

1.5.4 *New theories without new evidence*

In the history of phonological theory, new theories have often been preceded by new types of evidence. For example, the use of spectrography to examine the acoustic properties of speech led to Jakobson et al.'s (1952) acoustically defined feature system, and the use of Electromagnetic Midsagittal Articulometry (EMA) and X-ray microbeam technology led to Articulatory Phonology (e.g. Browman and Goldstein 1992). In other cases, the connection between new theories and new evidence is less overt. The claim that distinctive features are innate is one of these. Early feature theories did not claim innateness, but innateness is now a fairly standard assumption, and it is not clear what if any evidence brought about this shift.

In the early years of modern phonological theory, Trubetzkoy (e.g. 1939) and Jakobson stressed the importance of describing languages on their own terms. Jakobson (1942: 241) writes that "[t]he description of a system of values and the classification of its elements can be made only from that system's own perspective". Later, Jakobson takes more universalist views, but the evidence that leads to this conclusion is unclear. In part II of *Fundamentals of Language* (Jakobson and Halle 1956: 39), Jakobson claims that "[t]he study of invariances within the phonemic pattern of a given language must be supplemented by a search for universal invariances in the phonemic patterning of language in general". Further, Jakobson reports implicational relationships between

phonological distinctions, which are reportedly found in acquisition and in aphasia (Jakobson and Halle 1956: 38). While studying aphasia and acquisition would be expected to shed light on the structure and universality of distinctive features, none of the examples of aphasia given by Jakobson provides evidence for this. This work must be taken as an explication of the *predictions* of the theory, rather than empirical evidence in support of it. It is acknowledged more recently (by proponents as well as critics of his later universalist views regarding language acquisition) that Jakobson's model of language acquisition is based on his general theory of phonology rather than on actual language acquisition data (Menn 1980, Rice and Avery 1995). What is troubling about Jakobson's change of view is that unlike other developments, it is not accompanied by new evidence, but has nevertheless been widely accepted by phonologists who followed in his path.

Recent work on language acquisition has shown that children are highly individualistic in their order of acquisition of sounds and words (see Vihman 1993, 1996 for summaries). This is unexpected if a set of innate features is at the core of phonological acquisition. Research has shown that similarities between children acquiring language reflect the languages the children are learning, rather than universal tendencies (e.g. Ingram 1978, Pye et al. 1987, de Boysson-Bardies and Vihman 1991, Vihman 1996, Beckman, et al. 2003, Vihman and Croft 2004).

Another theoretical development which is not accompanied by any new evidence is the criterion that simplicity of representation should reflect the phonetic naturalness of a process, and that (according to Sagey 1986: 9–11) the phonological representation "should lead to explanation, where possible, of why the facts are as they are, and of why the representation is structured as it is". For example, the simplicity of the representation of a phonological pattern is argued to explain why it is more frequent than one with a more complex representation. The assumption that representations are explanatory in this way was not present in the bulk of early work on distinctive features (e.g. Jakobson 1942, Jakobson et al. 1952, Jakobson and Halle 1956, Chomsky and Halle 1968: chs. 1–8), but is assumed, apparently without any motivation, in many approaches to Feature Geometry. This has the effect of adding another dimension to the claim of distinctive feature universality (the need for the representation of one language to reflect markedness generalizations about language in general) without any argument for why such a representation is desirable, beyond aesthetic reasons (see e.g. Lass 1975 and Hume 2004b for counterarguments).

It is often assumed (e.g. Sagey 1986) that a representation that can be explained on the basis of factors such as vocal tract anatomy, acoustics, and

knowledge of the world is more highly valued than a representation which accounts for the same phonological facts arbitrarily. Not discussed, however, is the possibility that the phonological representation does *not* need to explain the non-occurrence of non-occurring segments, precisely because they do *not* occur. Sagey argues that segments such as doubly articulated palatal/velar stops should be unrepresentable because they are extremely difficult to produce as segments distinct from both palatal and velar stops. The hypothetical cognitive representation may be the last line of defense keeping doubly articulated palatal/velar stops out of human languages, but it is by no means the first. If no language ever develops them (for production-based reasons), then there is no need for the cognitive representation of phonological patterns to rule them out.

Sagey explicitly argues against including the Well-Formedness Condition (No Line-Crossing) in Universal Grammar, because it follows from knowledge about the world. This argument could also be leveled against phonetically grounded Feature Geometry as a whole, because the requirements it derives from are extralinguistic (physiological).

The role of features in acquisition and aphasia and the role of representations in reflecting the naturalness and frequency of phonological patterns are both relationships that are often treated as evidence for innate features. But these, like the ability of innate features to account for most if not all phonological patterns, are hypotheses. Acquisition and aphasia are the subject of much ongoing research, and the ability of feature theories to predict the frequency or possibility of sound patterns is challenged by the results of the crosslinguistic survey reported in later chapters.

1.5.5 *Dogs, fish, chickens, and humans*

Phonological features are sometimes treated as a uniquely human endowment which explains in part why humans acquire language but other animals do not. This is contrary to some of the early arguments for features, which involved evidence from the behavior of other animals to motivate key aspects of features. For example, in "The concept of phoneme", Jakobson (1942) treats distinctive features as a manifestation of the fundamental relationship between meaningful contrast and the ability to distinguish sounds. Jakobson observes that all native speakers of a given language can accurately perceive even the most minute phonetic differences as long as they perform a discriminative role, while foreigners, even professional linguists, often have great difficulty perceiving the same differences if they do not distinguish words in their own native languages. Jakobson's point is that there is a fundamental relationship between meaningful contrast and the ability to distinguish sounds, not that this has anything to do

with universality in the sense of Universal Grammar. Indeed, Jakobson notes that dogs and fish possess a similar faculty. The important distinction is between meaningful and non-meaningful differences, rather than between innately provided and non-innately provided differences. Jakobson gives examples of dogs being trained to recognize a particular pitch that signals the arrival of dog food, and to distinguish it from other, very similar pitches, as well as certain species of fish being trained to associate a certain acoustic signal with receiving food, and to associate another slightly different acoustic signal with "something nasty", so that the fish surface upon hearing one signal, hide upon hearing another, and ignore other signals. Jakobson (1942: 233) writes that the fish "recognize the signals according to their meanings, and only because of their meanings, because of a constant and mechanical association between signified and signifier".

Another parallel between the proposed nature of distinctive features and animal behavior is observed by Jakobson and Halle (1956: 26), involving relational rules. The opposition [compact] vs. [diffuse] (acoustic correlates of low vs. high vowels) characterizes the relation between [æ] and [e] and also the difference between [e] and [i]. Jakobson and Halle observe that the ability to understand such relations as instances of a single property is not unique to humans. They cite experiments in which chickens were trained to pick grain from a gray field, but not from a darker one, and when presented with a gray field and a *lighter* one, the chickens transferred the relation and picked grain only from the lighter field.

Much like the hypotheses involving aphasia, acquisition, and naturalness, the notion that features are part of the uniquely human ability to acquire language arose without direct evidence. Innate distinctive features are cognitive categories with built-in phonetic correlates. As shown by Jakobson, Halle, and others, cognitive category formation is shared with other members of the animal kingdom. Meanwhile, the phonetic correlates of features are not even shared by all human languages; spoken languages lack the correlates of signed language features, and vice versa. It is hard to imagine how a uniquely human capacity for language could involve innate distinctive features, when one aspect of supposedly innate features is too widespread and the other is too restricted. The use of innate distinctive features in phonology can be contrasted with syntax, which was the original motivation for Universal Grammar, and whose uniquely human innate component Hauser et al. (2002) have reduced to the operation of recursion.

1.5.6 *Innate features recapitulate independently observable facts*

Innate features have been used to account for a variety of observable facts about language. Often there are other explanations available for these facts,

and it may be the case that the feature theories are simply restating what is accounted for by other factors. Two ways in which this occurs are when synchronic formulations of phonological patterns appear to recapitulate historical changes, and when the feature organization which accounts for affinities between articulators appears to repeat explanations which are available simply from observing the physical relationships between them. For example, the model proposed in *SPE* accounts for a very wide range of sound patterns in modern English, often drawing on diachronic changes known to have occurred in the history of English. This approach has been criticized (e.g. by Pinker 1999: 100) as a recapitulation of the history of long-dead rules whose remnants can be memorized by modern speakers.

It is in large part because phonologists have had, over the past forty-odd years, an opportunity to build upon the groundwork laid by Chomsky and Halle that it is possible now to re-evaluate their claims. A critical re-evaluation of their assumptions about innate distinctive features would have seemed natural, but this is a path that mainstream phonological theory has not explored yet. Criticisms of the framework set forth in *SPE* are largely limited to Chomsky and Halle's choices of features and their organization, but do not address the basic assumption that there is a universal set of distinctive features. Chomsky and Halle's *assumption* that distinctive features are innate is treated in subsequent literature as if it were a *conclusion*.

While derivations often recapitulate historical changes, innate feature organization also encodes information that is independently observable. In motivating constituency among distinctive features, Clements (1985: 229) observes that at least four articulatory parameters show considerable independence from each other: (1) laryngeal configuration, (2) degree of nasal cavity stricture, (3) degree and type of oral cavity stricture, and (4) pairing of an active and a passive articulator. Oral tract configuration can be held constant while the state of the vocal folds or velum changes, and vice versa. However, within each category, it is difficult or impossible to vary one gesture while maintaining another.

There are external explanations for affinities between features and the properties they represent. For example, the claim (on the basis of patterning) that features such as [anterior] and [distributed] are dominated by the [coronal] node does not make particularly interesting predictions as long as these features are only used for coronal segments and are defined in terms of the coronal articulator. The organization recapitulates anatomical information which is built into the definitions of the features. A more compelling case for innate feature organization could be made on the basis of features which pattern in a certain way *in spite of* their phonetic definitions.

The incorporation of physiological information into formal phonology is taken to the extreme by Articulator Theories (Sagey 1986, Halle 1988, 1989, 1992, Halle et al. 2000), which directly incorporate anatomical adjacency as a criterion for feature organization. By incorporating anatomical adjacency rather than basing the model on phonological phenomena, Articulator Theories construct a model of the physiological facts which lead, via the phonologization of phonetic effects, to articulatorily driven phonological alternations. Drawing on physiological facts as a means of accounting for phonological patterns is not the same as including physiological facts in the representation of synchronic phonology. Including these facts in the representation is justified if it is motivated by observed phonological patterns that cannot be accounted for by other known factors.

Recent phonological theory has placed emphasis on explaining phonological patterns in terms of independent observations about phonetics and other factors. While this is a worthwhile pursuit, identifying these factors does not require repeating them in Universal Grammar. It may be true that these factors really are in the grammar, but motivating this requires more than just evidence that there is a pattern, because the pattern is already predicted by the external facts.

1.5.7 *Summary*

As seen in this chapter, there is substantial independent evidence calling innate features into question. The fact that quite a bit of what they account for may have other explanations anyway makes abandoning innateness quite reasonable. The formal model of the cognitive representation of phonology is often treated as if it is the only way to account for the nonexistence of unattested phonological patterns. This issue is particularly important when ruling out unattested phenomena compromises the ability of the formalism to capture some attested phenomena (such as unnatural classes), especially when there is no independent evidence that "unnatural" phenomena are treated any differently by speakers vis-à-vis common phenomena (see Buckley 2000, Onishi et al. 2002, and Peperkamp and Dupoux 2007 for additional discussion).

The notion of innate distinctive features clearly would not have remained popular for so long if there were not many correlations between phonological patterns and the phonetically grounded features that have been proposed to account for them. The question is this: When we study sound patterns, are we looking at something that innate features do that manifests itself in sounds, or are we looking at something sounds do that can be described with features?

The strongest position in support of innate features is one that perhaps has no proponents. This is what phonological patterns might be expected to be

like, given literal interpretation of the idea that features are the building blocks of phonological patterns (9).

(9) Innate features (strong position)
- All phonological patterns in spoken and signed languages can be reduced to operations on a small set of innate features.
- The role of phonetics in phonology can be reduced to the phonetic basis of distinctive features.
- A wide range of observations about phonological patterns can be attributed to facts about features themselves (e.g. their organization in the brain), with no interpretation in phonetics, language change, or anywhere else.

The weaker position in (10) is more widely held but harder to falsify. This position is informed by the observation that some phonological patterns are not easily interpretable as the manifestation of innate features. External factors are invoked to account for problem cases.

(10) Innate features (weak position)
- Most if not all *recurrent* phonological patterns in spoken and signed languages can be reduced to operations on a small set of innate features.
- The role of phonetics in phonology can *often* be reduced to the phonetic basis of distinctive features.
- Some observations about phonological patterns may be attributed to facts about features themselves (e.g. their organization in the brain), with no interpretation in phonetics, language change, or anywhere else.

The emergent features position in (11) dispenses with innate features as a means of accounting for observations about phonological patterns, and appeals directly to influences on phonological patterns.

(11) Emergent features
- Phonological patterns occurring with greater than chance frequency in spoken and signed languages can be accounted for in terms of external factors affecting them.
- The role of phonetics in phonology can be reduced to external factors (relating to vision, audition, articulation, etc.).
- No observations about phonological patterns may be attributed to facts about features themselves (e.g. their innate organization in the brain), with no interpretation in phonetics, language change, or anywhere else.

It should be clear that the strong version of the innate features position is not tenable. The purpose of this book is to motivate the emergent features position over the weak version of the innate features position. There are already many widely recognized external explanations for the existence, absence, or rarity of certain phenomena among the world's languages, and many of these are invoked in the weak version of the innate features approach. Two goals of emergent feature theory are (1) to show that when these external factors are taken seriously, there is nothing left for innate features to account for, and (2) to formalize the role of external factors in phonological patterns without including them in Universal Grammar or otherwise building them into the cognitive representation of phonology.

1.6 Original motivations for distinctive features

While there are many reasons to suspect that distinctive features are not innate, there are also many facts which distinctive features have been used successfully to account for. The approach in this book focuses on re-evaluating the insights of distinctive feature theory and recasting them in a framework that does not assume innateness, rather than discounting the contributions of innate feature theories to the study of phonology. This section summarizes some of the motivations for features and some of their typical properties.

1.6.1 *Motivations for features*

Features were proposed as a part of phonological theory long before they were argued to be innate. Early motivations for distinctive features focused on minimizing demands on memory and perception. Based on assumptions about a correlation between meaning and strain on perception and memory, Jakobson hypothesizes about a constraint on the number of phonological contrasts in a language:

> Differences which have differentiating value are, as we have seen, more accessible to perception and to memory than differences which have no value at all, but on the other hand differences between phonemes—since they lack particular meanings—strain perception and memory and necessarily require a great deal of them. We would expect, therefore, that the number of these primordial and unmotivated values would be relatively small for any given language. (1942: 235)

Because Jakobson assumes that the differences between phonemes, being "unmotivated", tax perception and memory, he argues that the number of oppositions should be minimized. If binary oppositions between phonemes

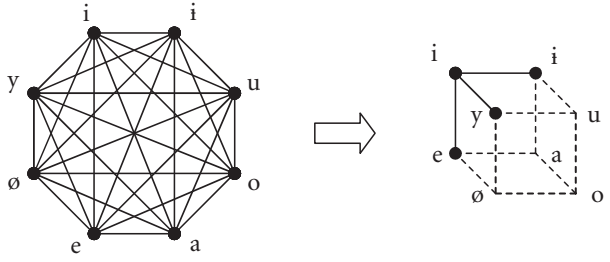


FIGURE 1.8 Reducing twenty-eight binary relations to three

are taken to be the "primordial" values, then twenty-eight (7+6+5+4+3+2+1) binary relations are necessary to characterize the eight vowels of Turkish. By introducing the notion of distinctive features, Jakobson reduces twenty-eight binary relations to three, as in Fig. 1.8.

For Jakobson, the argument for a minimal number of distinctive features in any given language is the same as the argument for the *existence* of distinctive features: It is assumed that primitives which have no inherent meaning are costly to perception and memory, and that their numbers in any given system are therefore minimized. Universality of distinctive features is limited to the claim that features in two languages which refer to the same acoustic feature are fundamentally the same. Thus, the feature [high] in Turkish is fundamentally the same as the feature [high] in Russian. In this sense, the set of possible phonological distinctive features is limited only by acoustic and articulatory phonetics, and the universality of the distinctive features (in spoken languages) is a direct consequence of the universality of the human vocal tract.

1.6.2 *Motivations for binarity*

The conclusion that distinctive features are binary was supported by Jakobson et al., on the basis of the observation that the distinctions between some pairs of words, such as *bill/pill* and *bill/dill*, can be characterized by a difference of one feature. Others are distinguished by more than one feature, such as pairs like *bill/fell*, which involve a duple distinction in initial segments and a minimal distinction in their middle segments. In essence, the fact that differences between words *can* be represented by a series of binary decisions is taken as evidence that this is actually how information is encoded in language. Jakobson et al. assert that Information Theory (e.g. Shannon and Weaver 1949) provides a sequence of binary selections as the most reasonable way to analyze communication, and that in the special case of language, this is not simply the best analysis to impose on the data, but how it is inherently

structured. While there is a continuous range of possible degrees of voicing and lip-rounding and other articulatory movements, only two polar points are picked out as distinctive features. Jakobson and colleagues argue that the dichotomous scale is the optimal code, and therefore there is no reason to suppose that speakers would use a more complicated system. However, they provide no evidence to show that this is limited to language rather than more general cognitive patterns of human beings (and perhaps also dogs, fish, and chickens). They report that binary relations are imprinted in children's early cognitive development (citing Wallon's 1945 study of gradual binary fissions in child development and Parsons and Bales' 1955 study of socialization), and note that almost all distinctive features are dichotomous at the articulatory and acoustic levels, and that applying the dichotomous scale makes the analysis of phonological patterns so clear that it must be inherent in language.

1.6.3 *Motivations for innateness*

The assumption of innate primitives in linguistic theory did not originate in the study of phonology. Chomsky's transformational grammar program, starting in the 1950s, crucially involved a universal, innate human language faculty containing formal and substantive linguistic universals. Formal universals correspond to the formalisms of linguistic theory, which are believed to be unlearnable, and therefore innate. The central component of linguistic competence in Chomsky's (e.g. 1957, 1965) program is syntactic, and so are the arguments for formal and substantive universals. *The Sound Pattern of English* (Chomsky and Halle 1968) represents a move to extend some of the formal universals of Chomsky's account of syntax, such as the transformational cycle, to the study of phonology. The claim set forth in Jakobson et al. (1952) that all the phonemes of the world's languages can be described in terms of twelve features is quite compatible with Chomsky's program. However, since many of the formal universals of syntax are no longer assumed (see e.g. Hauser et al. 2002), it is reasonable to reconsider some of the formal universals proposed for phonology on the basis of 1960s syntactic theory.

In contrast to previous accounts by Trubetzkoy (1939), Jakobson (1942), and Jakobson et al. (1952), Chomsky and Halle (1968) assume a cognitive, rather than physiological, basis for the universality of distinctive features. Distinctive features are provided by Universal Grammar, rather than determined by, for example, the universal vocal tract. While they acknowledge the role of the universal vocal tract in phonological patterns, Chomsky and Halle (1968: 14) propose that a phonetic representation is a feature matrix with rows corresponding to a small set of features and columns corresponding to

segments, and that "such representations are mentally constructed by the speaker and the hearer and underlie their actual performance in speaking and 'understanding' ".

According to Chomsky and Halle (1968: 164), distinctive features "must be determined absolutely, within general linguistic theory, and independently of the grammar of any particular language". This argument is based on the assumption that it is necessary for the functioning of their model and therefore necessary to the extent that their model works to explain English phonology. Because conditions such as the principle of the transformational cycle and the principles of organization of grammar do not seem to be learnable, these universals are hypothesized to be innate (Chomsky and Halle 1968: 43).

Phonology has never been central to the motivations for Universal Grammar, but many theories of phonology assume that primitives such as features are part of UG. Recently there have been a number of challenges to some of the more fundamental motivations for UG.[5] The more questionable the foundations of UG as well as the relationship between these foundations and phonology become, the more precarious the innate features position becomes.

This section examines the connections between arguments for Universal Grammar and the application of Universal Grammar in phonology. Many of the arguments for UG in other domains do not hold for phonology. For example, there is little evidence of a learnability problem in phonology (see Blevins 2004 for discussion).

Chomsky (1968: 124) considered the theory of universal phonetics to be much more fully established than the theory of universal semantics. This asymmetry could have been the result of the comparatively large amount of crosslinguistic work in phonetics and phonology (e.g. Trubetzkoy 1939, Jakobson et al. 1952, Chomsky and Halle 1968, and many others). Another possibility is that language sound systems seemed much more straightforwardly restricted in a way that could be attributed to Universal Grammar. Phonetics is well known to be constrained by physiology, and Jakobson found that a large number of sound systems can be described with a very small number of distinctive features. If the former is not treated as the cause of the latter, then Jakobson's distinctive features look like evidence for Universal Grammar. But if physiology is what constrains phonetics in such a way that it can be described with a small set of features, then neither observation is suggestive

[5] See Steels (1997) for a summary of some arguments against Universal Grammar.

of Universal Grammar. The fact that phonetics observations can be expressed with a small set of features has nothing to do with language-specific capabilities of the human brain, except perhaps that the human brain is usually in close proximity to the human vocal tract.

Jakobson observes that no language uses both labialization and velarization for distinguishing words, and that these could be variants of one abstract feature, and Chomsky (1968: 123) claims that such generalization can be proposed as laws of universal phonetics. Abstract generalizations are consistent with the notion of Universal Grammar as proposed for syntax, but the abstractness of the labialization/velarization feature is less clear when acoustics is considered in addition to articulatory phonetics. The acoustic correlate of both gestures is a lowering of F3, and the antagonistic relationship between labialization and velarization can be explained by the fact that they are perceptually indistinct. Invoking Universal Grammar is not necessary to explain Jakobson's observation.

The notion of universal phonetics can be stated in two ways: as a set of cognitive constraints in Universal Grammar, or as observations about the human vocal apparatus. The approach which would most strengthen Chomsky's (1968) position in general is the former, exemplified by the reference to an abstract feature responsible for labialization and velarization. But universal phonetics is often defined in the more trivial way, as in Chomsky and Halle (1968: 294–5), where it is the set of "phonetic properties that can in principle be controlled in speech". This definition is unassailable, but entails no cognitive explanation whatsoever for phonetic universals. Indeed, the phonetic motivation for Universal Grammar is extremely weak. Perhaps the most compelling case that can be made is that phonetics, like semantics, is part of the grammar, and that there is an implicit assumption that if syntax is rooted in Universal Grammar, the rest should be too. Most of the evidence for UG is not related to phonology, and phonology has more of a guilt-by-association status with respect to innateness.

1.7 Outline of the book

This chapter has raised a number of issues casting doubt on innate distinctive features. Emergent feature theory is developed as an alternative to innate features in Chapter 5. The three intervening chapters provide a little more background and some phonological evidence for emergent features. Chapter 2 reviews phonetic and psycholinguistic evidence that relates to distinctive features and/or their universality. Chapter 3 describes the methods of the crosslinguistic survey of phonologically active classes, and Chapter 4 gives a

first look at the results, focusing on "ambivalent" segments, which provide a means to separate the predictions of innate and emergent features. Chapters 5 and 6 present the results of the survey generally and in terms of three feature theories (Jakobson et al. 1952, *SPE*, and Unified Feature Theory). A general model of the emergence of linguistic structure is described in Chapter 8.