The Role of Features in Phonological Inventories

G. N. Clements

Laboratoire de Phonétique et Phonologie (UMR 7018)

CNRS / Sorbonne-Nouvelle, Paris, France

(final version, October 27, 2005)

Abstract. Phonological inventories are structured in terms of distinctive features, rather than finer-grained phonetic categories. Five feature-based principles are discussed and exemplified with respect to data drawn from a database containing 451 phoneme inventories. By *Feature Bounding*, features place an upper bound on the number of potentially contrastive categories in a language. By *Feature Economy*, features tend to be combined maximally. By *Marked Feature Avoidance*, certain feature values tend to be avoided. By *Robustness*, highly-valued feature contrasts tend to be employed before less highly-valued contrasts. By *Phonological Enhancement*, marked feature values may be introduced to reinforce weak perceptual contrasts. These principles interact to predict broad properties of sound systems, such as symmetry and the tendency of sounds to be dispersed in auditory space. Further phonetically-based principles fine-tune the realization of phonological categories at the phonetic level. It is suggested that these general properties of sound systems may find an explanation in the nature of early language acquisition.

1. Introduction

As linguists have long noted, not just any set of consonants and vowels can make up a phonological inventory. A central finding of the earliest work in phonology was that speech sound systems are structured in terms of recurrent elementary components known as features (e.g. Trubetzkoy 1969 [1939], Martinet 1955, Hockett 1955).[1]

More recently, however, the study of inventory structure has been subject to some neglect. The intensive effort devoted within the generative tradition to discovering the architecture of rule systems (or more recently, constraint systems) has not been matched by similar efforts in the area of phonological inventories. This may be due to the belief that inventories have no existence independent of the lexicon and that generalizations regarding their structure are external to the grammar as such. Although this was not the position of Chomsky & Halle, their remarks on the subject (1968, chapter 9) focused almost exclusively on markedness and barely touched on such further issues as economy and feature hierarchy.

For these and other reasons, in recent years the nature of inventory structure has been more vigorously debated among phoneticians than among phonologists (see e.g. Maddieson 1984). Two general approaches have emerged, based on the role they assign to features. In one, which we might term a *feature-mediated* theory of inventory structure, sound systems are viewed as constrained by the fact that speech is perceived and produced in terms of distinctive features. In this approach, features are viewed as biologically grounded in that they correspond to articulatory regions that have relatively stable, distinctive acoustic properties. Inventory-based generalizations are typically formulated over natural classes of sounds as defined by features. This approach has been exemplified notably in the work of Kenneth N. Stevens and his colleagues.

In an alternative approach, features play little or no role. In this approach, which we might term a *direct-access* theory of phonological explanation, generalizations about speech sound inventories -- including surface-phonemic inventories in the classical sense -- refer directly to the finer-grained categories provided by phonetic theory. These categories may include:

- the minute, and in some cases infinitely divisible articulatory categories postulated by descriptive phonetics (e.g. Maddieson 1984, Ladefoged & Maddieson 1996)
- the auditory and articulatory variables employed in a model which views phonological inventories as emerging from the interplay of auditory dispersion and articulatory ease (e.g. Lindblom 1996, 1992, Lindblom & Maddieson 1988)
- the articulator sets and parameter settings employed in gesture-based phonetics, which models phonological systems in terms of gestures and their interactions (e.g. Browman & Goldstein 1989, 1992, 2000)

These trends of research has been salutary in bringing to light many respects in which phonological patterning is shaped by constraints imposed by the medium of speech itself, and has introduced a necessary corrective to the "overly formal" approach to inventory structure (Chomsky & Halle 1968: 400) taken by classical generative phonology. However, by neglecting features, these approaches appear to make phonetic explanation incommensurable with phonological structure. They raise the following question: if features are the principal categories in terms of which phonological systems are structured, why should they be irrelevant to universals of phonological inventories?

While there has been valuable work on inventory structure by generative phonologists, this work has tended to emphasize descriptive formalisms over system-level principles. Work in mainstream Optimality Theory has reinforced the neglect of inventory structure, due to the fact that constraint systems usually evaluate individual forms rather than system-wide generalizations. More recently, however, some OT-oriented phonologists have proposed to incorporate system-level principles into the theory (Boersma 1997, Flemming 2002). Even in this work a bias toward phonetic reductionism can be observed, to the point that a contemporary linguist can maintain that the study of contrast “does not require a restrictive inventory of distinctive features" but that "phonological representation can include the entire sea of predictable or freely varying phonetic detail".

This paper reviews a range of evidence showing that distinctive features play a central role in structuring inventories of contrastive speech sounds. It examines a number of general principles that appear to be most straightforwardly stated if we take features as arguments. These principles heavily constrain the shape of preferred sound inventories, and make strong and testable predictions regarding the trends we may expect to find as we examine as yet undescribed languages.

The discussion proceeds as follows. Section 2 outlines the general view of features that will be assumed here. Section 3 presents the data base used in this study. The next sections review five feature-based principles that significantly constrain the structure of sound systems: Feature Bounding (section 4), Feature Economy (section 5), Marked Feature Avoidance(section 6), Robustness (section 7), and Phonological Enhancement (section 8). Section 9 applies these principles to illustrative cases, and section 10 discusses some implications of these results for phonological theory.

2. Features: their nature and cognitive status

We first review some fundamental aspects of feature theory, stressing its phonetic and cognitive grounding.

Features have been defined from the very beginning in concrete physical terms, though linguists have hesitated between auditory, acoustic and articulatory definitions. Trubetzkoy suggested that though acoustics is basic to the definition of speech units, articulatory definitions must still be used for most practical purposes as “acoustic terminology unfortunately is still very sparse” (1939: 92). The acoustic study of speech was greatly advanced by the work of Jakobson and his collaborators in the 1950s, who were able to draw upon technological advances in the study of speech such as the use of the Sonograph (sound spectrograph). They believed, however, that the most relevant definitions of speech units lay in the perceptual and auditory domains:

> The closer we are in our investigation to the destination of the message (i.e. its perception by the receiver), the more accurately can we gage the information conveyed by its sound shape. This determines the operational hierarchy of levels of decreasing pertinence: perceptual, aural, acoustical and articulatory (the latter carrying no direct information to the receiver). The systematic exploration of the first two of these levels belongs to the future and is an urgent duty. (Jakobson, Fant & Halle 1952: 12)

Though the auditory and perceptual domains are somewhat better understood a half-century later, there is still nothing approaching a consensus on how the properties of speech sounds are to be defined in these terms. Subsequent work on feature theory, inspired in part by the motor theory of speech perception (Liberman et al. 1967), has given new prominence to articulatory definitions (Chomsky & Halle 1968), though it is usually agreed that feature have acoustic correlates as well (Lieberman 1970, Halle 1983).

More recently, a new integration of articulatory, acoustic and perceptual approaches to feature definition has been achieved within the context of quantal theory (Stevens 1972, 1989). This approach is based on the observation that there are continuous articulatory regions within which moderate changes in the positioning of an articulator have essentially negligible acoustic and perceptual effects, while at the boundaries between such regions small articulatory movements have significant effects. The stable regions typically define distinctive features. Thus, for example, as the tongue blade is retracted from the dental to the alveopalatal region in the production of fricatives, the acoustic spectrum undergoes abrupt, perceptually salient changes as the constriction passes through the unstable region corresponding to the boundary separating anterior sounds such as [s] from posterior sounds such as [ʃ]. This boundary, then, separates [+anterior] sounds from [-anterior] sounds. Within the class of [+anterior] sounds or [-anterior] sounds, in contrast, acoustic differences are too small (all else being equal) to be used in defining distinctive contrasts. In this approach, features are grounded in objectively observable discontinuities between production and perception.[2]

Quantal theory provides a new basis for understanding fundamental aspects of speech processing. One of the more robust results emanating from psycholinguistic work in recent years is that speech is processed in categorical terms. In discrimination and identification tasks, adult speakers of a language have been found to identify and distinguish speech stimuli more accurately *cross* the phoneme boundaries of their language than within them (Liberman et al. 1957, Repp 1984, Harnad 1987). In contrast, perception of non-native contrasts falling within these boundaries is rather poor. Though other studies have shown that perception of non-native contrasts can be improved by explicit training or prolonged exposure to a second language, adult speech perception remains relatively inflexible in comparison to the plasticity shown by very young learners (Pallier et al. 1997). The phonetic dimensions for which categorical perception has been confirmed, including voicing and major place of articulation contrasts, correspond closely to those used by quantally-defined distinctive features.[3]

There is much evidence that categorical perception is present even at birth, and that infants abstract phoneme-like categories from irrelevant "noise" in the signal (such as differences among speakers). For example, full-term new-born babies (aged from 2 to 6 days) have been shown to discriminate between the syllables *pa* and *ta*, regardless of whether the stimuli were spoken by the same speaker or by any of four different speakers (Dehaene & Pena 2001). Moreover, categorical perception in infants is not restricted to the features of the mother tongue. English babies have been found to discriminate non-English place of articulation contrasts that exist in Hindi (Werker et al. 1981) and non-English vowel contrasts that exist in French (Trehub 1976), while infants 6 to 8 months old discriminate non-English contrasts found in Hindi and the Interior Salish language Thompson (Werker & Tees 1984).

Such findings suggest that the ability to distinguish major phonetic categories – typically corresponding to quantally-defined features -- exists in early infancy. In later infancy, the ability to discriminate speech sounds becomes "fine-tuned" to the categories of the native language. By 6 months of age, infants have established prototypes for the vowels used in their language (Kuhl et al. 1992) and start to loose sensitivity to nonnative vowels (Polka & Werker 1994). By 12 months, they seem have lost the capacity to discriminate nonnative consonantal contrasts which can be assigned to a single native category (Werker & Tees, 1984). After that, the capacity to perceive foreign contrasts remains generally poor, although it is still fairly good if the foreign contrasts can be assimilated to native contrasts, or if they are perceived as non-speech sounds (Best et al. 1988, 2001).

It appears, then, that there is a continuity in speech processing from early infancy to adulthood and that the predisposition for categorical perception and normalization shown by infants carries over to a large extent to adult speech processing. If this is so, one might naturally expect phonological inventories to be structured in terms of the same broad feature-based categories that appear in the course of language acquisition. This hypothesis will be explored in the rest of this study.

3. Data and Method

This study examines cross-linguistic trends in the structure of sound inventories. Evidence is drawn primarily from the expanded UPSID data base as described in Maddieson & Precoda (1989). This data base presents several advantages. First, it contains phoneme inventories drawn from 451 languages, representing about 6-7% of the world's languages, according to current estimates. Secondly, it was constructed by selecting just one language from each moderately distant genetic grouping, assuring a minimum of genetic balance. Third, its electronic format facilitates rapid searches, eliminating the need for the laborious and time-consuming scrutiny of printed materials such as Maddieson (1984). Fourthly, since the data base is publicly available from the UCLA Phonetics Laboratory where it was compiled, results obtained from it can be independently verified by others.

However, even the best available data base is necessarily imperfect. A number of problems in the UPSID data base have been discussed by Basbøll (1985), Maddieson (1991), Simpson (1999) and Clements (2003a), among others. These include:

- an inevitable skewing toward the properties of larger genetic units (e.g. the Niger-Congo unit is represented by 55 languages, Basque by only one);
- the heterogeneity of the primary sources and disagreements in analyses;
- the inclusion of some allophonic details but not others (e.g. the dental vs. alveolar stop distinction is registered even when noncontrastive, while the apical vs. laminal stop distinction is rarely noted);
- the occasionally inconsistent choice of basic allophones for each phoneme;
- the presence of a fair number of coding errors

To a considerable extent, these problems are alleviated by the sheer size of the sample. Generalizations supported at a high level of significance by large numbers of genetically diverse languages are unlikely to be far off the mark, and most generalizations discussed in the present study are of this type. However, caution must be taken in interpreting results, especially when there is a likelihood of error or oversight in the primary sources. Many of the sources used in compiling UPSID are less than fully reliable, and in such cases other sources should be consulted as well.

"Inventories" such as one finds in UPSID are abstractions over sounds that are contrastive in a language and typically include consonants that may appear in different positions in the syllable and word. Most consonants of a language, however, can appear word-initially, and consonants that appear elsewhere are usually a subset of these. A consonant inventory usually approximates an inventory of word onsets, and the sounds selected as the basic allophone or variant of a consonant phoneme in UPSID are typically those that appear in strong positions such as the onset. The rationale for this choice is that consonants that appear elsewhere can often be regarded as reduced or lenited realizations of this basic variant. (For fuller discussion of the criteria used in selecting basic allophones in UPSID see Maddieson 1984, 62-3; 1991, 196.)

For the purposes of the ongoing research of which this study is a part, the phoneme systems of UPSID have been recompiled in terms of phonological features. A fairly conservative feature system has been used, using widely familiar features similar to those proposed in Halle & Clements (1983), to which the articulator features [labial] and [dorsal] of Sagey (1990) have been added. While further revisions of this system have been suggested (see e.g. Halle 1992, Clements & Hume 1995), for present purposes these more familiar features will be adequate.

Most results reported in this study have been tested for statistical significance with the chi square test, which is typically used to find out whether two independent characteristics are associated in such a way that high frequencies of one tend to be coupled with high frequencies of the other. The .01 level of probability is taken as criterial. See Clements (2003a) for further discussion of the statistical method.

4. Feature Bounding

Let us now review a number of feature-based principles that appear to govern the structure of speech sound inventories. The first is one which I shall call *Feature Bounding*. This principle involves two claims. One is that features set an upper limit on how many *sounds* a language may have. More exactly, given a set of *n* features, a language may have at most 2^n distinctive sounds. For example, a language using three features may have up to eight sounds (2^3), one using four features may have up to sixteen sounds (2^4), and so forth. No more are possible.[4]

Secondly, features also set an upper limit on the number of *contrasts* that may appear in a language. The possible contrasts (C) in a language are a function of the total number of its sounds (S) and is given by the expression $C = (S * (S-1)) / 2$. Given that the maximum number of possible sounds is 2^n, the maximum number of contrasts for a system with *n* features is therefore

$(2^n * (2^n - 1))/2$. By this calculation, for example, a language with 2 features may have up to 4 sounds and 6 contrasts.

Coronal consonants provide an illustration. Feature theory proposes two features distinguishing "major place of articulation" in coronal sounds, defined in terms of the location and form of the front-of-the-tongue constriction along the midline of the oral cavity. These are anterior/posterior and distributed/nondistributed. These two features define four classes of sounds, as shown below:[5]

(1)

	apico-anterior	lamino-anterior	retroflex	postalveolar/palatal
posterior	-	-	+	+
distributed	-	+	-	+

"Major place" as just defined is a broader notion than "place of articulation" in traditional phonetic theory, which recognizes many more place categories within the coronal region. By providing a larger set of distinctions, phonetic theory admits a greater number of potential contrasts. Table 1 makes this point by comparing the maximum number of sounds and contrasts predicted by a feature theory making use of the four categories shown in (1) and those predicted by a phonetic theory recognizing the seven categories "apico-dental", "apico-alveolar", "lamino-dental", "lamino-alveolar", "palato-alveolar", "retroflex", and "palatal".

	Max. no. sounds	Max. no. contrasts
a. feature theory	4	6
b. traditional phonetic theory	7	21

Table 1. Maximum number of distinct coronal sounds and maximal number of coronal contrasts predicted by a feature system recognizing 4 coronal categories (row a) and a phonetic theory recognizing 7 coronal categories (row b).

As row (a) shows, a feature theory providing just 2 feature and 4 major coronal categories predicts up to 6 potential contrasts. In contrast, as row (b) shows, a traditional phonetic theory recognizing 7 coronal categories predicts up to 21 potential contrasts, more than three times the number predicted by feature theory.

The predictions of feature theory appear to be correct in this case. I have elsewhere reported plausible attestations for all 6 predicted contrasts, not only among simple stops but among strident stops (affricates) as well (i.e. Clements 1999). Examples of the six predicted contrasts for simple plosives are listed in (2), with illustrative languages drawn from Ladefoged & Maddieson (1996) shown at the right.

(2) Contrasts among coronal plosives:

contrast:	example:	found in e.g.:
apical anterior *vs*. nonapical anterior	apical *t* *vs*. nonapical *t*	Temne
apical anterior *vs*. apical posterior	apical *t* *vs*. retroflex *ʈ*	Yanyuwa
apical anterior *vs*. nonapical posterior	apical *t* *vs*. palatal *c*	Arrernte
nonapical anterior *vs*. apical posterior	nonapical *t* *vs*. retroflex *ʈ*	Toda
nonapical anterior *vs*. nonapical posterior	nonapical *t* *vs*. palatal *c*	Ngwo
apical posterior *vs*. nonapical posterior	retroflex *ʈ* *vs*. palatal *c*	Sindhi

Most strikingly, *no other primary coronal contrasts* were found in either plosives or affricates in a survey of several hundred languages. (The sample comprised all 451 languages of the expanded UPSID data-base and several other languages known for their rich coronal inventories.) In particular, no reliable example was found of a minimal contrast, unaccompanied by any other feature difference, between dental and alveolar stops or between palato-alveolar and palatal stops such as are predicted by the traditional IPA categories.[6]

Proposed contrasts beyond the six predicted by the features [±posterior] and [±distributed] have not been substantiated. I have elsewhere discussed several alleged cases of this type (Clements 1999) and have shown that they do not require additional coronal categories. Beyond the cases discussed there, Ladefoged & Maddieson (1996, 42) cite two allegedly minimal contrasts between apico-dental and apico-alveolar sounds which prove, on closer examination, to be accompanied by other featural differences. First, Albanian is said to contrast apical dental and apical alveolar laterals. Such sounds cannot be distinguished by the features assumed here since both are [-posterior] and [-distributed]. However, a study of the source, Bothorel (1969-70), shows that the distinction between the two apical sounds transcribed *l* and *ll* is accompanied by a further distinction involving the position of the tongue body. As Bothorel describes it (p. 135), the essential difference between the two laterals comes from the lowering of the entire body of the tongue for *ll*, with a consequential retraction of the tongue root, narrowing of the pharyngeal passage, and opening of the lateral passages, a configuration distinct from the conic form and gradual lowering we find in *l*. Examination of his x-ray figures confirms that *ll* is indeed strongly backed with respect to *l*, a difference which can be expressed by the secondary articulation features [dorsal] or [pharyngeal]. Second, Ladefoged & Maddieson state that in many Khoisan languages such as !Xóõ, some speakers have an apical dental contact for the dental click /|/ and an apical alveolar contact for the alveolar click /!/ (1996, 42). However, as they point out elsewhere in the same work (p. 257-9), these sounds have a prominent difference that is far more regular across speakers: /!/ is produced with an abrupt release while /|/ is realized with an affricated release, making the first plosive-like and the second affricate-like. This difference parallels similar differences between non-click stop types in these languages and can be described with the feature [±strident].

It appears, then, that the features [±posterior] and [±distributed] successfully characterize the set of primary coronal contrasts that is actually attested across languages. It is not obvious how a feature-free account of phonetic structure could predict this set of contrasts.[7] It might be thought, perhaps, that phonetic theory could exclude unattested contrasts on the basis of a principle of "sufficient dispersion" according to which sounds must meet a minimum criterion of auditory distinctness in order to contrast in a phonemic system. In such a view, the fact that few if any languages have minimal contrasts between apical dental and apical alveolar stops would be explained by the observation that these two sounds are auditorily very similar to each other. It appears, however, that the contrasts overgenerated by traditional phonetic categories -- such as dental vs. alveolar stops -- *are just those that cannot be described in terms of phonological features*. Once we eliminate such contrasts by assigning them to the appropriate feature categories, we obtain the attested number of distinctive sounds and contrasts. There is no need to appeal to a special theory of sufficient dispersion for this purpose since auditory dispersion is built into feature theory itself, through its requirement that features be specified for quantally distinct attributes.

5. Feature Economy

Feature Economy is the tendency to maximize feature combinations (see Clements 2003a,b, after sources in de Groot 1931, Martinet 1955, 1968). This principle can be observed in most speech sound inventories, regardless of size. Let us consider, by way of illustration, the surface-distinctive consonants of a standard variety of English as shown in (3), focusing attention on the sounds in the box:

(3)

p^h	t^h		$tʃ^h$	k^h	
b	d		dʒ	g	
f	θ	s	ʃ		
v	ð	z	ʒ		
m	n			ŋ	
w	l, r		y		h

It can be seen that voicing cross-classifies stops and fricatives to double the number of obstruents; this feature is used with maximum efficiency in the obstruent subsystem. Though the feature [+continuant] is used with less efficiency (since English lacks the fricatives /x/ and /ɣ/), it nevertheless creates two full fricative series. The feature [+nasal] creates nasals stops at three places of articulation. At the other extreme, the feature [+lateral] is used with minimal efficiency, as it only distinguishes the pair /l/ and /r/.[8]

Though the vast majority of languages exhibit Feature Economy to some degree, no language makes use of all theoretically possible feature combinations. For example, English fails to combine nasality with obstruence to created a series of nasal fricatives. As observed by Martinet (1955), such gaps often correspond to functionally inefficient feature combinations and tend to be widely avoided across languages. Thus, nasality is inefficient in fricatives as it is difficult to achieve the air pressure buildup required in the production of fricative noise while allowing air to pass through the nasal cavity. (We return to a discussion of markedness considerations in the next section.)

Feature Economy can be quantified in terms of a measure called its *economy index* (Clements 2003a). Given a system using F features to characterize S sounds, its economy index E can be expressed, to a first approximation, by the equation in (4).

(4) $E = S/F$

The higher the value of E, the greater the economy. For example, if we make use of the following 9 features to distinguish the 24 English consonants:[9]

(5) labial, dorsal, glottal, posterior, continuant, voiced, strident, nasal, lateral

the economy index of the English consonant system is 24/9, or 2.7.

Feature Economy can be defined as the tendency to maximize E. This goal can be accomplished either by

- increasing the number of sounds, but not features, or
- decreasing the number of features, but not sounds

(or eventually both). Both strategies are exemplified in phonological systems. First, increasing of the number of sounds while holding the number of features constant is reflected in historical changes which create new phonemes from existing features (Martinet 1955). A familiar example is the historical creation of a new series of [+nasal] vowels in French through the historical deletion of syllable-final [+nasal] consonants. Second, decreasing the number of features while holding the number of sounds constant is reflected in the frequent historical elimination of "isolated" sounds that do not fall into regular patterns of correlation with other sounds; after elimination of such sounds, the feature that previously characterized them becomes redundant. An example of the latter tendency can be cited from two stages of Zulu, as described in Clements (2003a).

(6)

stage 1:			stage 2:		
p’	t’	k’	p’	t’	k’
p^h	t^h	k^h	p^h	t^h	k^h
b	d	g	p	t	k
		k	b		g
ɓ					

In stage 1, reflecting the usage of a century ago, we find two isolated stops, the implosive *ɓ* and the plain voiceless *k*, both of which are the sole members of their series. Through a subsequent evolution whose end product is shown in stage 2, the voiced stops devoiced and the two isolated sounds subsequently shifted into a single voiced series, as shown in the last row. Stage 2 differs from stage 1 in its increased economy, since the feature which previously distinguished the implosive from its plain voiced counterpart has been eliminated.[10]

Feature Economy must be distinguished from the more familiar criterion of *parsimony,* which requires the use of the fewest units possible in any given analysis. Like economy, parsimony favors reducing the number of features, but it also militates against increasing the number of sounds, and thus fails to predict historical trends in which existing features recombine to yield new sounds (as in the evolution of French nasal vowels). Feature Economy is also different from *symmetry*. Symmetry, like Feature Economy, requires the number of gaps in a system to be minimized, but would not usually be viewed as penalizing a 3x3 system such as that in (7).

(7)

p	t	k
b	d	g
f	s	x

However, this system is not fully economical, as it is missing a voiced fricative series [v z ɣ] combining the existing features [+voiced] and [+continuant]. Feature Economy predicts that a voiced fricative series should tend be present if the three series shown in (7) are present as well. (This prediction is correct, as we shall see below.)

How may such predictions be tested? Earlier work examined the predictions of Feature Economy in the historical domain (Martinet 1955). In a recent publication I have outlined a method for testing this principle at the synchronic level and have applied it to the phoneme systems of the expanded UPSID data base (Clements 2003a). One prediction of Feature Economy is *Mutual Attraction*, which can be stated as follows:

(8) A given speech sound will have a higher than expected frequency in inventories in which all of its features are distinctively present in other sounds.

For example, according to this prediction, a voiced labial fricative V should be more frequent in systems having some other labial consonant such as B, F, or N, some other voiced obstruent such as B, D, or Z, and some other fricative such as F, X, or Z. This prediction can be tested by constructing a 2x2 contingency table as shown in Table 2. (Upper-case letters here and below denote general feature-defined classes rather than specific phonetic values.)

		some other labial *and* some other voiced obstruent *and* some other fricative?		
		yes	no	*total*
V?	yes	136 (114)	11 (33)	147
	no	214 (236)	90 (68)	304
	total	350	101	451

Table 2. Observed frequencies of voiced labial fricatives (V) across UPSID languages, according to whether the language does ("yes" column) or does not ("no" column) have another labial, another voiced obstruent, and another fricative. Expected frequencies are shown in parentheses.

Reading the cells from left to right and top to bottom, this table presents the number of languages which:

- have V together with some other labial, some other voiced obstruent, and some other fricative (136);
- have V, but lack a labial, a voiced obstruent, or a fricative (11);
- lack V, but have another labial, another voiced obstruent, and another fricative (214);
- lack V, and also lack a labial, a voiced obstruent, or a fricative (90).

Parenthesized numbers show the values that would be statistically expected under the assumption that phonemes combine randomly, contrary to the predictions of Feature Economy.[11] For example, the number of languages that would be statistically expected to have V together with some other labial, some other voiced obstruent and some other fricative on the assumption of random combination is 114, which is much lower than the 136 that we actually observe. The difference between observed and expected values in this case is highly significant under chi square testing ($\chi^2 = 27.902$. $p < .0001$) and confirms prediction (8): V is indeed significantly more frequent in systems in which its distinctive features are independently present in other sounds. This trend reveals Feature Economy at work; the features [labial], [+voiced] and [+continuant], once present in a system, tend to recombine to form other sounds.

Table 3 shows the result of testing 18 pairs of stop consonants for economy effects by this method. The consonants in each pair share all manner features, but differ in place. Feature Economy predicts that if a feature combination appears at one place of articulation, it should tend to appear at other places of articulation as well. For example, if a system contains a labial implosive we expect it to contain a coronal implosive, and vice-versa. In this table, the symbols P, T, K stand for any voiceless labial, coronal, or dorsal stop, respectively, and B, D, G for any voiced labial, coronal, or dorsal stop. Diacritics indicate manner features as explained in the legend.

P^- vs. T^-	P^h vs. T^h	P’ vs. T’
P^- vs. K^-	P^h vs. K^h	P’ vs. K’
T^- vs. K^-	T^h vs. K^h	T’ vs. K’
B vs. D	B^h vs. D^h	$B^<$ vs. $D^<$
B vs. G	B^h vs. G^h	$B^<$ vs. $G^<$
D vs. G	D^h vs. G^h	$D^<$ vs. $G^<$

Table 3. Comparisons among pairs of stops sharing all manner features, but differing in place of articulation. Symbols: P^- T^- K^- = plain voiceless stops, P^h T^h K^h = voiceless aspirated stops, P’ T’ K’ = ejective stops, B D G = voiced unaspirated stops, B^h D^h G^h = voiced aspirated stops, and $B^<$ $D^<$ $G^<$ = implosive stops.

All these comparisons test positive at a very high level of significance ($p<.0001$). That is, languages having one member of each pair tend overwhelmingly to have the other.

Let us return to the consonant inventory in (7). Under Feature Economy, though not symmetry, we expect that a system containing the sounds /p t k/, /b d g/, and /f s x/ will also have the voiced fricatives /v z ɣ/, maximizing the use of [+voiced] and [+continuant]. This prediction is confirmed by chi square testing. Voiced labial fricatives such as /v/ are considerably more frequent than expected in sample languages having all three of the sounds /p/, /b/, and /f/. Fully analogous results hold for/ /z/ and /ɣ/. These trends are significant at the .0001 level in all cases.

In Clements (2003a) I have discussed other examples illustrating a variety of Feature Economy effects. These results show that sound systems are strongly structured by Feature Economy, a principle defined, as its name implies, in terms of features. Here again, phonetically-based accounts do not appear to perform as well. One such alternative, Gesture Economy (Maddieson 1995), can be shown to be less adequate in predicting phoneme inventory structure, though it may correctly predict the tendency toward articulatory uniformity at the level of phonetic implementation.

To summarize, this section has reviewed evidence for a principle of Feature Economy according to which sound systems tend to maximize the use of a small number of features. The economy of a system can be quantified in terms of an economy index, E. Feature Economy places pressure on systems to increase this index, either by reducing the number of features (holding the number of sounds constant) or by increasing the number of sounds (holding the number of features constant). This principle is confirmed by an examination of statistical trends in the UPSID data base.

6. Marked Feature Avoidance

Markedness, as it applies to inventories, can be understood as the tendency to avoid certain widely disfavored feature values – *marked* values (see Trubetzkoy 1969 [1939], Jakobson 1968 [1941], Greenberg 1966, Chomsky & Halle 1968, Kean 1980, Calabrese 1994, and others, as well as Rice 2000 and Gurevich 2001 for critical overviews). This section proposes a new approach to the study of inventory markedness based on a principle of Marked Feature Avoidance, which replaces the traditional notion of implicational universal.

6.1. Markedness and Feature Economy

We first consider how Markedness interacts with Feature Economy. In its strongest form, Feature Economy predicts that all languages make use of all possible combinations of all distinctive features in constructing their phoneme inventories. However, no language comes even close to achieving this goal. This is due to at least two factors. First, if some feature values are redundant – that is, predictable from other features -- they are available to serve as secondary cues to the presence of other, distinctive features, especially in contexts where the primary cues to the distinctive feature are weak or absent. Second, some feature values – the marked values -- appear less suited to speech communication than others for articulatory and perceptual reasons. If marked values are *under*used, not only is the redundancy of the system increased, aiding processing, but the articulatory and perceptual complexity of the system is substantially reduced.

The English consonant system shown in (3) provides a good example of feature under-utilization. If this system used all possible combinations of its 9 consonant features, it would have 2^9 or 512 consonants (including a number of highly unusual sounds) instead of the 24 relatively common sounds it actually contains. The English system is thus heavily constrained by markedness.

It is perhaps less obvious that Feature Economy also counteracts Markedness. However, this is also a strong effect. Once a feature value is present in a system, Feature Economy creates pressure for it to be used again, even if that value is marked. An example is provided by voiced fricatives, which bear the marked values [+voiced] and [+continuant]. These values are marked in fricatives presumably because they operate to counteract the buildup of supraglottal air pressure required to produce the turbulence noise characteristic of obstruents (e.g. Stevens 1983). For this (and perhaps other) reasons, voiced fricatives are absent in roughly half the world's languages. However, due to the effect of Feature Economy, voiced frictives are far more likely to occur in languages that have other voiced fricatives. This is demonstrated in (9), which shows the percentage of UPSID languages having voiced labial, coronal, and dorsal fricatives 1) overall, 2) in languages having one or more other voiced fricatives, and 3) in languages lacking other voiced fricatives.

(9) occurrence of voiced fricatives in UPSID

		numbers	%
[labial]	overall:	147/451	32.6
	in languages having other voiced fricatives:	111/184	60.3
	in languages lacking other voiced fricatives:	36/267	13.5
[coronal]	overall:	174/451	38.6
	in languages having other voiced fricatives:	129/175	73.7
	in languages lacking other voiced fricatives:	45/276	16.3
[dorsal]	overall:	70/451	15.5
	in languages having other voiced fricatives:	62/212	29.2
	in languages lacking other voiced fricatives:	8/239	3.3

Thus, languages are much more likely to have two or more voiced fricatives than to have just one. Again, English with its four voiced fricatives illustrates this trend.

The tendency for Feature Economy to override Markedness can be shown in another way. Pericliev & Valdés-Pérez (2002) have documented the occurrence of what they call *idiosyncratic segments* in UPSID, that is, segments occurring in just one language. They find that if a language possesses several such segments, they strongly tend to share features that bind them together as a class. Thus, in 44 of the 53 languages having more than one idiosyncratic segment, all such segments share features, while in only 4 languages were all such segments unrelated. A study of their data shows that in the great majority of cases, the shared features involve marked values, as in the case of Arrernte with its six unique nasally released stops or Shilha with its seven unique pharyngealized consonants. The likelihood of such configurations arising from random distribution is, of course, extremely small.

Thus Feature Economy and Markedness operate antagonistically, Feature Economy tending to expand the size of an inventory and Markedness to contract it. Sound systems represent varying degrees of compromise between these two forces.

6.2. Which value of a feature is marked?

Markedness statements, like other types of phonological statements, usually refer to natural classes of sounds, taking the form of general statements such as “nasal vowels are marked with respect to oral vowels” or “fricatives are marked with respect to stops". Specific statements such as "/ã/ is marked with respect to /a/" or "/θ/ is marked with respect to /t/" are usually instances of broader generalization such as these. Markedness is therefore a property of classes of sounds, as defined by marked feature values. Segments can be said to be marked just to the extent that they bear marked feature values.

Given this fact, markedness theory requires a criterion for determining the marked value of any feature, within and across languages. Linguists have generally taken one of two approaches to this question. One has sought to define markedness in terms of the substantive conditions that underlie the human capacity for speech production and speech processing. From this point of view, marked features are sometimes said to be those that are relatively difficult to implement, or that lack salient acoustic properties. A severe problem for this approach, however, is that the various fields embraced by phonetics – acoustics, physiology, neurology, aerodynamics, auditory perception, etc., -- constitute a number of interacting complex systems, no one of which explains all aspects of speech, and which taken together often make conflicting predictions. There does not yet appear to be any general, overarching theory that predicts unambiguously, for any given feature, what its marked value must be.

A second approach to defining markedness is based on frequency, or likelihood of occurrence (e.g. Kean 1980). As pointed out by Greenberg (1966) and others, markedness tends to be reflected in frequency differences at many levels; thus, marked segments tend to be less frequent in the lexicon, in texts, in early stages of language acquisition, and in adult sound inventories, and tend to show fewer contextual variants. Moreover, it is well established that human are sensitive to frequency distributions in the data to which they are exposed from infancy onward (see e.g. Jusczyk et al. 1994, Maye et al,. 2002, Anderson et al. 2003). Though a frequency-based criterion cannot explain why some segments are more frequent than others, it has the advantage of relating markedness to quantitative trends that are observable by language learners and which plausibly form part of the input in the construction of the target grammar (Pierrehumbert 2003).[12]

For such reasons, it may be preferable to interpret markedness as an effect of frequency. The less frequent value of a feature is, by virtue of its relative unexpectedness in discourse, the more salient one, and for this reason may exhibit the special properties often associated with marked values, such as their tendency to engage in assimilation, to trigger dissimilation, to “float” in the absence of segmental support, and so forth.

The markedness criterion that will be adopted here is stated in (10):

(10) A feature value is marked if it is absent in some languages, otherwise it is unmarked

A feature is counted as absent in a language if it does not occur in the primary allophone of some contrastive sound. For example, [+nasal] sounds are absent in this sense in some languages (including Quileute, Lushootseed, Pirahã, and Rotokas), while all known languages have [-nasal] sounds. It follows that [+nasal] is the marked value of the oral/nasal distinction. Features that are marked in terms of criterion (10) usually satisfy other markedness criteria as well. In the case of [+nasal], for example, we find that:

- speech sound inventories tend to contain fewer [+nasal] sounds than [-nasal] sounds
- most lexicons contain fewer [+nasal] sounds than [-nasal] sounds
- [+nasal] sounds tend to have lower text frequencies in most languages
- in any inventory, [+nasal] stops usually imply corresponding [-nasal] stops
- [+nasal] is more likely to spread from one segment to another
- [+nasal] sounds often neutralize to the nearest [-nasal] sound in non-assimilatory contexts, while the reverse is not true

In evaluating a given feature by criterion (10), only classes of sounds in which the feature is potentially distinctive are considered. For example, since all vowels are redundantly [+sonorant] and [+continuant], vowels are not counted in evaluating these features. However, as these features are potentially distinctive in consonants, the marked value can be determined by examining consonants. Thus we observe that some languages (e.g. Pirahã and Maxakalí) lack [+sonorant] consonants while all have obstruents, and that others (e.g. Auca, Dera, Angaatiha, and Ekari) lack [+continuant] consonants though all have stops. It follows from (10) that [+sonorant] and [+continuant] are the marked values of these features.

Examples of marked feature values following this criterion are given in Table 4.

all languages have:	*some lack:*	*marked feature values:*
obstruents	sonorant consonants	[+sonorant]
oral stops	oral continuants	[+continuant]
oral sounds	nasal sounds	[+nasal]
nonstrident sounds	strident sounds	[+strident]
anterior coronal sounds	posterior coronal sounds	[+posterior]
nonlateral sounds	lateral sounds	[+lateral]
unaspirated sounds	aspirated sounds	[spread glottis]
nonglottalized sounds	glottalized sounds	[constricted glottis]
unrounded sounds	rounded sounds	[+rounded]
nonhigh vowels	high vowels	[+high]
nonlow vowels	low vowels	[+low]
central and back vowels	front vowels	[+front]

Table 4. Marked feature values according to criterion (10). Data source: UPSID.

Not all features can be assigned marked values by criterion (10). For example, since all languages have both vowels and consonants, neither value of [±consonantal] can be considered marked by this criterion. The UPSID data base does not give sufficient information to determine whether [±distributed] or [±ATR] can be assigned marked values. (For further discussion of these features see Clements 2001 and Casali 2003, respectively.)

Criterion (10) extends straightforwardly to one-valued features like [spread glottis]. As is shown in the table, [spread glottis] is marked; since not all languages make use of it; the unmarked term is simply the absence of this features. Similar considerations argue for the marked status of most articulator features. Thus, as consonant features, [labial] and [dorsal] are marked, since some languages (e.g. Wichita) have no primary labial consonants while others (e.g. Vanimo) have no primary dorsal consonants. In contrast, all known languages appear to have coronal consonants. (Hawaiian, often cited for its absence of /t/, has /n/ and /l/.) Criterion (10) therefore identifies [coronal] as unmarked.[13]

The feature [±voiced] requires special discussion. While most languages have voiceless obstruents, voiced obstruents are lacking in many languages. This would suggest that [+voiced] is the marked value of this feature.[14] However, UPSID includes four languages, all spoken in

Australia, in which *all* obstruents are classified as voiced and are transcribed with the symbols *b, d, g*, etc. It would follow from (10) that neither value of this feature is marked.

A closer look at the facts, however, shows that the "voiced" obstruents in these languages are realized as voiceless in some circumstances. For example, in Mbabaram, according to a more recent description, "the allophony is as follows: (i) stops are normally voiceless in initial position, in final position, and in medial position after *y*; (ii) they are voiced after a nasal; (iii) they alternate between voiced and voiceless in medial position between vowels, or after l, r, or ɽ" (Dixon 1991: 355). Many linguists would feel quite comfortable transcribing such sounds with the symbols *p, t, k*. However, the fact that they are often voiced between vowels shows that they are not quite the same as the voiceless stops of languages like French or Spanish. UPSID also records Dyirbal and Yidiny as having only voiced stops. Again, however, other sources show that they have regular voiceless variants. For Dyirbal *b*, Dixon observes: "The voiced allophone [b] is almost invariably heard between vowels (which are always voiced sounds); at the beginning of a word [p] is often heard, varying freely with [b] in this position. That is, [diban], [tiban] are the two commonest pronunciations of 'stone'. [dipan], [tipan] are heard much less often; but they are, unhesitatingly, taken as instances of the same word" (1980, 127). For Yidiny, Dixon states: "Stops are almost always voiced. Partly voiced allophones are sometimes encountered word-initially ..." (Dixon 1977, 32). Generalizing over Australian languages, Yallop observes: "The plosives of aboriginal languages may be pronounced sometimes as voiced sounds (*b, d*, etc.) and sometimes as voiceless sounds (*p, t*, etc.) – but the voiced and voiceless counterparts are either freely interchangeable or in complementary distribution" (Yallop 1982, 56). The validity of this claim is confirmed by the study of some of the more detailed descriptions of individual Australian languages.[15]

Such realization patterns suggests that the "voiced" stops in question are produced with a laryngeal configuration in which the vocal cords are adducted as for modal voicing but neither tensed nor laxed. Phonetic studies show that whether the vocal cords vibrate in such a state is a function of the interacting articulatory and aerodynamic factors that regulate pressure drop across the glottis (see e.g. Westbury 1983, Stevens 1998). Given this fact, the variably voiced stops of many Australian languages might be better analyzed as *lacking* any specification for voicing. The presence (or absence) of voicing would then be determined not by phonological features but by the phonetic context. Such an analysis can be easily expressed, for example, in the feature theory of Halle & Stevens (1971, 1991), which recognizes three categories of unaspirated, non-glottalized stops:

- voiced stops, bearing [+slack vocal cords]
- voiceless stops, bearing [+stiff vocal cords]
- a third category of "intermediate" stops, bearing neither feature

Stops that vary between voiced and voiceless, freely or according to the phonetic context, are quite naturally viewed as representing the third category of stops.[16] If voicing is reanalyzed in this way, the problem raised by Australian languages disappears. In these languages, stops do not bear a feature [+voiced], but lack specifications for both [stiff vocal cords] and [slack vocal cords]. In this analysis, [stiff vocal cords] and [slack vocal cords] are *both* marked features, since not all languages make use of them.

6.3. Marked Feature Avoidance

On the basis of these observations, we may state the principle of *Marked Feature Avoidance* as follows (< = “is less than”):

(11) Within any class of sounds in which a given feature F is potentially distinctive, the number of sounds bearing marked values of F < the number bearing unmarked values of F

This principle claims that languages tend to avoid marked feature values, regardless of the class of sounds they occur in. It predicts, for example, that [+nasal] sounds will be less frequent than [-nasal] sounds in the classes of vowels, liquids, sonorants, etc. Like Feature Economy, this principle represents a force rather than a strict law, and can be expected to have exceptions. We examine several such exceptions in section 8, where we will see that they arise from the interaction with a competing principle of Enhancement. The claim, then, is that (11) forms one of a number of interacting principles that together govern the structure of sound inventories.

Some specific claims that follow from (11) are shown below:

(12)	a. nasal vowels < oral vowels	*marked feature*:	[+nasal]
	b. fricatives < stops	*marked feature*:	[+continuant]
	c. sonorant consonants < obstruents	*marked feature*:	[+sonorant]

These statements bear a superficial resemblance to what are commonly called implicational universals. Such universals are usually given in the form “the presence of M implies the presence of U”, or more simply “M $\supset$ U”, in which M is a marked term and U the corresponding unmarked term. Thus, we find statements such as those in (13):

(13) a. nasal vowels $\supset$ oral vowels
b. fricatives $\supset$ stops
c. sonorant consonants $\supset$ obstruents

We may observe that the truth of these statements follow strictly from the truth of those in (12). Thus, for example, if the number of nasal vowels in a languages is always less than the number of oral vowels, in accordance with (12a), it can never be the case that a language will have nasal vowels without oral vowels, violating (13a).

However, while the statements in (13) follow from those in (12), the reverse is not true. Languages can violate the statements in (12) without violating those in (13). Consequently, those in (12) make the stronger claim. To see this, let us compare the vowel systems in (14):

(14)	system A	system B	system C
	i u		
	e o	õ	õ
	a ã	a ã	ã

Both (12a) and (13a) exclude system C, which has only nasal vowels. Both admit system A, in which nasal vowels form a proper subset of oral vowels. However, they differ in their predictions regarding system B. This system is admitted by the implicational universal (13a) since the presence of the single oral vowel satisfies the implied term U ("oral vowels"). However, it is excluded by Marked Feature Avoidance, since the number of nasal vowels exceeds the number of nasal vowels.

Now it is a common observation that the number of nasal vowels in a vowel system is typically a proper subset of the number of oral vowels; systems like A are common, while those like B are apparently unattested (see e.g. Williamson 1973, Maddieson 1984). It follows that even if we adopt the universal implication (13a), we would *still* need principle (12a) to express this widespread trend. Once we admit (12a), however, (13a) becomes superfluous, as it is a logical consequence of (12). Thus the implicational universal (13a) may be dispensed with. A similar line of reasoning applies to other implications such as those of (13b,c).

It may be objected that implicational universals such as those in (13) are still required because they are exceptionless by definition, while those in (12) state trends, and are therefore not always true. For example, the existence of a single language with more nasal vowels than oral vowels would not weaken the statistical validity of (12a), but would disprove the implicational universal "nasal vowels $\supset$ oral vowels". The claim, then, is that "absolute" implications such as those in (13) are still required to state exceptionless patterns. This objection is ill founded, however. "Absolute" implications such as those in (13) are exceptionless not because of any inherent link between the implying term and the implied term, but simply because the implied term occurs in all languages. Why is this? It will be recalled that a material implication of the form $P \supset Q$ is true if Q is true, no matter whether P is true or false. It follows that the statement "nasal vowels $\supset$ oral vowels" is true by virtue of the fact that all languages have oral vowels. Though the statement is formally true, it is misleading in suggesting a causal connection between nasal vowels and oral vowel: oral vowels are not present in a system because nasal vowels are present, but because (for quite different reasons) all languages have oral vowels. The statement "labial clicks $\supset$ oral vowels" is equally true, for example, yet there is patently no causal connection between labial clicks and vowels.[17]

Marked Feature Avoidance is a more powerful tool than implicational universals in still another sense. (15) presents further predictions of this principle in which the implied term is *not* found in all languages.

(15) a. in the class of uvular obstruents, fricatives < stops
marked feature: [+continuant]
b. in the class of voiceless stops, ejectives < pulmonic stops
marked feature: [constricted glottis]

In these cases there are no corresponding implicational universals:

- uvular fricatives do not universally imply uvular stops: several languages, including Pashto, Armenian, Mandarin, Spanish, and Basque, have uvular fricatives without uvular stops
- voiceless ejective stops do not universally imply voiceless plain (unaspirated pulmonic) stops: at least one language, Berta, has a weakly ejective series without a voiceless plain stop series (Triulzi et al. 1976)

Nevertheless, these statements represent valid trends, as we see from the data in (16):

(16)

sound type	no. lgs.	no. sounds	average per lg.
a. uvular fricatives	49	100	2.0
uvular stops	69	169	2.4
b. ejective stops	68	248	3.6
plain stops	435	2200	5.1

Overall, the marked member of each comparison is less frequent than the unmarked member, as predicted by Marked Feature Avoidance. Here, again, implicational universals have nothing to say.

There is an important class of exceptions to Marked Feature Avoidance, which upon closer examination proves to be explained by its interaction with Feature Economy. It will be recalled from the earlier discussion that Feature Economy tends to override Markedness. In its strongest form it predicts, contrary to Marked Feature Avoidance, that the marked members of a class should be equal in number to the unmarked members of a class. Now this is often true. For example, a fair number of languages have equal numbers of oral and nasal vowels. In a survey of 141 representative African vowel systems, I found that 45 had nasal vowels, and that of these, 7 or 15.6% had equal numbers of nasal and oral vowels. The following system, for example is found in Ikwere, spoken in Nigeria (Clements & Osu, in press)

(17)	i			u	ũ
	ɪ			ʊ	ʊ̃
	e	ẽ		o	õ
	ɛ	ɛ̃		ɔ	ɔ̃
			a ã		

Similarly, many languages, including English (3), have equal numbers of voiced and voiceless fricatives, or voiced and voiceless stops, and so forth. In all such cases, Feature Economy overrides Marked Feature Avoidance. The combined prediction of these two principles, when taken together, is that the number of marked sounds may be equal to but will never *exceed* the number of corresponding unmarked sounds in a class.

While this prediction is very largely true across languages, there are nevertheless some cases in which marked sounds do outnumber the corresponding unmarked sounds. For example, Archi, a Lezghian language, has more uvular fricatives (twelve) than stops (ten) (Colarusso 2004). In other cases, a marked sound may have no unmarked counterpart at all. As noted in the discussion of (16) above, for instance, several languages have the uvular fricative /χ/ without a corresponding uvular stop /q/. It must be emphasized, however, that most such cases represent minority patterns rather than general trends. [18]

6.4. Marked segment types appear in larger inventories

A final prediction of Marked Feature Avoidance is stated in (18):

(18) The average number of sounds in languages containing a marked term M is greater than the average number of sounds in languages containing its unmarked counterpart U.

This is because languages having a marked class of sounds M generally contain an equal or larger number of sounds of the unmarked class U, while the reverse is not true. An example will make this clear. Consider typical systems containing an array of voiced and voiceless stops, with or without corresponding fricatives, as shown in (19). (Upper-case letters designate general feature-defined categories rather than specific sounds.)

(19)

system A			system B			system C			system D		
P	T	K	P	T	K	P	T	K	P	T	K
B	D	G	B	D	G	B	D	G	B	D	G
			F	S		F	S				
						V	Z		V	Z	

The first three systems are well-formed in terms of Marked Feature Avoidance. System D, however, violates Marked Feature Avoidance as it has more voiced fricatives (2) than voiceless ones (0). The prediction, then, is that systems of this type should tend not to occur. This proves to be strongly and significantly true; while many UPSID languages have both series of fricatives, as in system C, or just voiceless fricatives alone, as in system B, only seven have voiced fricatives alone (χ^2=22.377, p<.0001). The conclusion is that voiced fricatives will tend to be admissable in a system only if voiceless fricatives are already in place.

We expect, then, that systems with voiced fricatives will tend to have more consonants overall (since they must also have voiceless fricatives, as in system C) than will systems with voiceless fricatives (which need not have voiced fricatives, as in system B). This expectation is confirmed in UPSID, as shown in (20).

(20)

sound:	total lgs.	average no. of consonants
voiced fricatives	221	22.3
voiceless fricatives	406	20.4

Languages with voiced fricatives have, on average, about two consonants more than languages with voiceless fricatives. (Compare systems B and C in (19).)

This prediction extends to more complex cases involving sounds with highly marked features. As an example, let us examine the distribution of plain dorsal stops K, labialized dorsal stops K^W, ejective dorsal stops K’, and labialized-ejective dorsal stops K^W’. The three latter categories bear the marked features of glottal constriction and labialization. Marked Feature Avoidance predicts that the labialized-ejectives K^W’ should tend to be present only in inventories that also have plain labialized K^W and simple ejectives K’, and that these in turn should tend to appear only in the still larger set of inventories that also have plain non-ejective K. As a consequence, K^{W}' should appear, on average, in languages with the largest consonant inventories, K^{W} and K' in languages with somewhat smaller inventories, and K in languages with the smallest inventories.

The numbers corresponding to each of these cases in UPSID are shown in (21).

(21)

	sound:	marked feature values	total lgs.	average no. of consonants
a.	K^{W}'	2	23	35.8
b.	K'	1	68	29.0
	K^{W}	1	69	26.4
c.	K	0	450	19.7

Our expectation is confirmed. The 23 languages containing K^{W}', bearing both marked features [constricted glottis] and [+rounded], have an average number of 35.8 consonants; languages with K' and K^{W}, bearing just one of these features, have an average number of 29.0 and 26.4 consonants, respectively; and languages containing K, bearing neither of these features, have an average of 19.7 consonants (equal to the average number of consonants in UPSID languages overall). In such examples, we find a positive correlation between the degree of markedness of a segment and the average size of the inventories containing it.

6.5. Summary

This section has made the following main points:

- marked feature values can be defined as those that are not present in all languages
- inventories show a tendency to avoid marked feature values
- by Marked Feature Avoidance, this tendency holds in all classes of sounds
- Marked Feature Avoidance interacts with other principles such as Feature Economy (see also the discussion of Enhancement in section 8)

The ultimate question of why some features values are more frequent (and hence more marked) than others has not been addressed, much less solved, in this discussion, and remains a question for further research.

7. Robustness

A further principle structuring sound inventories is one that will be called *Robustness*. This principle holds that there is a universal hierarchy of features such that languages draw upon higher-ranked feature in the hierarchy before drawing upon lower-ranked features in constituting their inventories. (For this and related ideas see Jakobson 1968 [1941], Jakobson & Halle 1956, Chomsky & Halle 1968: 409-410, Stevens & Keyser 1989, Dinnsen 1992, Calabrese 1994, 1995, Lang & Ohala 1996.)

The Robustness principle addresses a significant gap in the theory developed so far. Sound inventories do not typically consist of only vowels, or only fricatives, or only labial sounds. Instead they typically draw their members from a wide variety of feature dimensions, including at least three major places of articulation and several manner categories (Maddieson 1984). Feature Economy does not predict this distribution. Markedness theory goes some way toward accounting for it, but does not explain why sound systems are not uniformly skewed toward unmarked categories. Why do we find no languages whose consonant inventories include only coronals, or only voiceless stops? Why are there no languages with only central vowels?

The answer seems to be that languages prefer to draw their sounds from a highly differentiated set of sounds which are distinguished along many acoustic/articulatory parameters. A language having only coronal consonants would fail to benefit from the rich set of auditory contrasts that become available once labial, dorsal, and laryngeal consonants are introduced. Similarly, a language with only voiceless stops would not fully exploit the resonance properties of the vocal tract, and a language with only central vowels would not make full use of the frequency spectrum. The point of drawing sounds from many well-differentiated phonetic dimensions is that the members of a system built up in this way are highly individualized and distinct from one another.[19]

Robustness theory, then, is based on the observation that some contrasts are highly favored in sound systems, others less favored, and others disfavored. Contrasts can be arranged in terms of an approximate hierarchy according to the degree to which they are favored; contrasts high in the hierarchy tend to be present in most languages, while those lower in the hierarchy are present in fewer languages. It follows that contrasts lower on the list will tend to be present an inventory only if contrasts higher on the list are also present.

The earlier literature often failed to distinguish Robustness from Markedness. The basic difference is that Markedness is a property of feature *values* while Robustness is a property of feature-based *contrasts*. Some examples of more vs. less robust contrasts are given in (22):

(22)

more robust:	*less robust:*
sonorant vs. obstruent	apical vs. nonapical
labial vs. coronal vs. dorsal	central vs. lateral
nasal vs. oral	aspirated vs. nonaspirated
stop vs. continuant	glottalized vs. nonglottalized
voiced vs. voiceless	implosive vs. explosive

Like Markedness, Robustness is ultimately rooted in phonetic and functional factors. The most robust features, as a class, have the property of ensuring a high degree of dispersion of a "core" set of speech sounds. Robust features are, in general, those that maximize salience and economy at a low articulatory cost. They tend to permit one sound to be easily distinguished from another, even in rapid speech and under conditions of noise, and they are often mastered fairly early in the process of language production, one criterion of articulatory ease. Another factor that supports Robustness is economy -- the ability of a feature to combine freely with other features. Thus, for example, [±continuant] combines with all places of articulation, and [labial] with all manners of articulation. In contrast, many less robust features, such as [±strident], [±lateral], [±distributed], and [spread glottis], combine less easily, or not at all, with certain other features.[20]

However, as in the case of Markedness, the phonetic basis of Robustness is still poorly understood. Stevens & Keyser (1989) have suggested that "primary" features -- in our terms, those that stand at the top of the Robustness Scale -- provide a stronger auditory response than others. They state that "the three primary features are especially closely tied to fundamental capabilities of the auditory system for processing temporal and spectral aspects of sound" (p. 87). Thus, for example, [-continuant] obstruents (i.e. stops) are distinguished from [+continuant] obstruents (i.e. fricatives) by an abrupt decrease in amplitude over a wide range of frequencies, contrasting with neighboring intervals of high amplitude; contrasts of this type appear to trigger enhanced auditory responses in the peripheral auditory system (e.g. Delgutte & King 1984). It is far from clear, however, that all primary features can be characterized in this way. For example, clicks also involve rapid change in amplitude, yet these sounds are rare across languages. As Stevens & Keyser (1989) note, "We cannot, at this point, quantify the saliency of individual features in terms of auditory response mechanisms" (p. 85). Given these problems, a frequency-based diagnostic of Robustness will be adopted here, as in the case of Markedness.

(23) lists the commonest consonant contrasts in UPSID by order of frequency. These fall into four provisional groups, within which contrasts have roughly similar frequencies. Each contrast is illustrated by a typical pair, whose least frequent member is shown first.[21] Values in

percent indicate how many languages have consonants representing each contrast (for example, 91.6% of UPSID languages have contrasts between fricatives and stops). The features defining each contrast are given at the right.

(23)

	example	% (UPSID)	feature
a. dorsal vs. coronal obstruent	K / T	99.6	[dorsal], [coronal]
sonorant vs. obstruent	N / T	98.9	[±sonorant]
labial vs. coronal obstruent	P / T	98.7	[labial], [coronal]
labial vs. dorsal obstruent	P / K	98.7	[labial], [dorsal]
labial vs. coronal sonorant	M / N	98.0	[labial], [coronal]
b. continuant vs. noncontinuant sonorant	J / N	93.8	[±continuant]
continuant vs. noncontinuant obstruent	S / T	91.6	[±continuant]
posterior vs. anterior sonorant	J / L	89.6	[±posterior]
c. voiced vs. voiceless obstruent	D / T	83.4	[±voiced]
oral vs. nasal noncontinuant sonorant	L / N	80.7	[±nasal]
d. posterior vs. anterior obstruent	TŠ / T	77.6	[±posterior]
glottal vs. non-glottal consonant	H / T	74.5	[glottal]

Let us briefly consider contrasts based on [±strident], not included in this list. Two of the contrasts involving obstruents could also have been characterized in terms of [±strident]: the continuant vs. stop contrast (e.g. S/T) and the anterior vs. posterior contrast (e.g. TŠ/T). In both cases, one member of the contrast is typically a strident sound. However, [±continuant] and [±posterior] have been chosen as the basis of the contrast, for two reasons. First, these features are used more frequently than [±strident] across languages to defined minimal contrasts, even if we accept that [±strident] distinguishes simple stops and affricates (see Clements 1999, Kim 2001, Kehrein 2002). Thus, 404 UPSID languages contrast coronal stops and fricatives (involving [±continuant]), 212 contrast anterior and posterior coronal stops (involving [±posterior]), and 187 contrast anterior and posterior fricatives ([±posterior] again). In contrast, only 178 contrast sibilant and non-sibilant coronal stops ([±strident]), and just 80 contrast sibilant and non-sibilant coronal fricatives ([±strident]). Of the three features, then, [±strident] is the least often used to define minimal contrasts. Secondly, [±strident] can be understood as a feature that enhances the acoustic properties of continuants (fricatives) and posterior sounds (such as palato-alveolars), in the sense of Stevens, Keyser & Kawasaki (1986). It enhances fricatives by increasing the amplitude of the frication noise at higher frequencies, and posterior obstruents by making their characteristic lower-frequency noise component in the region of the third formant more audible (see further discussion in the next section). If one feature enhances another, the latter is considered more basic.

It seems appropriate, then, to regard [±continuant] and [±posterior] as the basis of these contrasts, even when [+strident] is also present in one member of the contrast. In the case of [±continuant], this choice is also supported by Feature Economy. While [±continuant] can be generalized to all oral places of articulation (labial, coronal, dorsal), [±strident] is restricted to coronal places (Sagey 1990), and thus tends to contribute less to the overall economy of a system. One would therefore expect [±continuant] to be the more basic feature even in segments in which both features are present.

A language having just the favored contrasts in (23) would typically include the consonants and glides shown in (24), assuming common realizations:

(24)					
	P	T	TŠ	K	
	B	D		G	
		S			
	M	N			
	W	L,R	J		H

Let us call this the Basic Consonant Inventory. It contains the fifteen commonest consonant types in UPSID, as defined by the feature contrasts in (23). Their frequencies are shown below.

(25)	sound type	feature definition	percent	typical phonetic values
	T	obs vl ant cor stop	98.2	t
	K	obs vl dor stop	97.8	k q
	N	son nas ant cor stop	95.6	n
	M	son nas lab stop	94.7	m
	P	obs vl lab stop	90.2	p
	S	obs vl cor cont	88.9	s ʃ
	J	son oral pos cor cont	85.1	j
	L	son oral cor stop	81.4	l ʎ
	W	son oral lab cont	80.3	w
	H	glot	74.5	h ʔ
	B	obs vd lab stop	71.4	b
	R	son oral ant cor cont	71.0	r
	D	obs vd ant cor stop	70.3	d

TŠ	obs vl pos cor stop	66.5	tʃ c
G	obs vd dor stop	63.2	g

These consonant types are defined by the features [±sonorant], [labial], [dorsal], [coronal], [±nasal], [±continuant], [±consonantal], [±voiced], [±posterior], and [glottal], as shown in the second column of (25). These features are presumably among the most robust for consonants.[22]

Based on these observations, a partial Robustness Scale is proposed in (26) for the more important consonant features, with the most robust features placed at the top. Features within each of the first four groups are unordered (ordering within group (e) remains to be determined).

(26) Robustness Scale for consonant features

a. [±sonorant]
   [labial]
   [coronal]
   [dorsal]
b. [±continuant]
   [±posterior]
c. [±voiced]
   [±nasal]
d. [glottal]
e. others

(26) expands the two-point scale proposed by Stevens & Keyser (1989). According to (26), the most robust features are [±sonorant] and the three major place features, [labial], [dorsal], and [coronal] (group a). These features are made use of in the great majority of languages. The remaining features are drawn upon with decreasing frequency as we descend the scale:

- [±continuant] and [±posterior] further expand the set of places and manners of articulation;
- [±voiced] and [±nasal] introduce the laryngeal and nasal dimensions, respectively;
- [glottal] is used here to designate sounds using a primary glottal articulator, namely H-sounds and glottal stops.

Remaining features such as [±strident], [±distributed], [±lateral], [spread glottis], and [constricted glottis] are less widely drawn upon and tend to be used only if the higher-ranked ones are also exploited.

A language making full use of features (a-d) in their most favored segmental contexts will have all members of the Basic Consonant Inventory, as shown in (27).

(27)

	P	T	TŠ	K	B	D	G	S	H	M	N	L	R	W	J
sonorant	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
labial	+				+					+				+	
coronal		+	+			+		+			+	+	+		+
dorsal				+			+							+	
continuant	-	-	-	-	-	-	-	+	+	-	-	-	+	+	+
posterior		-	+			-		-			-	-	-		+
voiced	-	-	-	-	+	+	+	-	-						
nasal	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-
glottal									+						

On the basis of the Robustness Scale (26) we may formulate the *Robustness Principle* as follows:

(28) In any class of sounds in which two features are potentially distinctive, minimal contrasts involving the lower-ranked feature will be present only if minimal contrasts involving the higher-ranked feature are also present.

It predicts, for example, that minimal contrasts involving [±strident] (group (26e)) will be present only if minimal contrasts involving [±voiced] (group (26c)) are present, that those involving [±voiced] will be present only if those involving [±continuant] (group (26b)), are present, and so forth.

As an illustration, let us consider the three partial consonant systems in (29). All contain a plain coronal stop T and a coronal affricate, either anterior TS or posterior TŠ.

(29)

system A				system B				system C			
P	T	TS	K	P	T	TŠ	K	P	T	TS	K
B	D	DZ	G								

Systems A and B are consistent with the Robustness Principle (28). System A has minimal contrasts involving [±strident] (TS/T, DZ/D) and [±voiced] (P/B, T/D, etc.); since higher-ranked

[±voiced] is present, lower-ranked [±strident] is also allowed. System B contains no minimal contrast involving either feature. The contrast T/TŠ is not minimal, as it involves two features, [±posterior] and [±strident]. Of these, the first is more robust, according to the Robustness Scale (26). Accordingly, the fundamental contrast in this case involves [±posterior], with [±strident] serving as a redundant feature of enhancement (see the next section for further discussion). System C, however, violates the Robustness Principle (28), as it contains a minimal contrast involving lower-ranked [±strident], namely T/TS, while lacking a minimal contrast involving higher-ranked [±voiced].

The Robustness Principle would, of course, be too strong if it were taken to be an exceptionless law. As in the case of Feature Economy and Marked Feature Avoidance, it is proposed as one of a number of interacting forces, and is intended to express significant trends, rather than laws. As we would expect, then, some languages and language groups have exceptions to it. For example, Spanish with its /r/ vs. /l/ contrast makes use of [±lateral] while abjuring the higher-ranked [glottal]. Similarly, as shown in (30), the Mexican language Zoque with its /t/ vs. /ts/ contrast draws upon [±strident] in its native obstruent system while not exploiting the higher-ranked [±voiced] (Wonderly 1951-2).

(30)						
	p	t	ts	tʃ	c	k
			s	ʃ		

Interestingly enough, though the feature [±voiced] is not distinctive in the native lexicon of Zoque it is distinctive in its substantial stock of Spanish-derived loanwords, and also appears pervasively in the native phonology by virtue of a regular process which voices stops after nasals (Wonderly 1946, 1951-2). These redundantly voiced stops appear to be phonetically identical to the phonemically voiced stops of the Spanish-derived lexicon. Such observations suggest that feature contrasts that are passed over in phoneme inventories may have a high potential for subsequent incorporation as a result of contact or internal change.

Although we do, therefore, find a certain amount of crosslinguistic variation, such variation is within narrow limits. For example, we find no languages which select all (or even most) of their features from the very bottom of the scale. Moreover, exceptions, when they exist, tend to be concentrated at the lower end of the scale, containing features that are easily dispensed with; while very few languages skip high-ranked features such as [±sonorant] or [labial], low-ranked features like [±lateral] or [±distributed] can be passed over with relative ease.

Robustness interacts strongly with Feature Economy and Marked Feature Avoidance. First, as a result of Feature Economy, even though less robust features tend to be less frequent across languages, once they are present in a system they tend to generalize to other sounds, creating further contrasts. For example, though [spread glottis] is a lower-ranked feature, languages with distinctively aspirated stops tend to have many of them. A further interaction is that Feature Economy favors features that combine maximally with others, reinforcing Robustness. Thus the robust feature [±sonorant], which cross-classifies all oral-cavity consonants, combines more readily with other sounds than the less robust feature [±strident], which cross-classifies only [coronal] sounds. The fact that [±strident], [±distributed], and [±lateral] are lower-ranked features is due, in part, to the fact that they combine poorly or not at all with noncoronal sounds.

Second, Marked Feature Avoidance limits the full generality with which even the most robust features are used. For example, labial and dorsal fricatives are absent in most languages, even though [labial], [dorsal] and [±continuant] are highly robust, and Feature Economy favors their maximal use. The Basic Inventory (24) reveals these gaps. The disfavored sounds are just those that cumulate marked feature values, [+continuant] together with [dorsal] or [labial]. Other common gaps, also illustrated in the Basic Inventory, are the following:

- dorsal nasals are missing: [+nasal] is disfavored with [dorsal] obstruents
- posterior fricatives are missing: [+posterior] is disfavored with [+continuant] obstruents
- posterior nasals are missing: [+posterior] is disfavored with [-continuant] sonorants

In these cases, too, the missing combinations cumulate marked feature values. It is often the case that if an expected higher-ranked contrast is missing, it is missing in one or more marked categories. This follows directly from Marked Feature Avoidance (11). In consequence, contextual conditions such as "[dorsal] only in [-continuant] obstruents" have an independent explanation, and do not have to be built into the Robustness Scale itself.[23]

In summary, this section has outlined a principle of Robustness which states that languages tend to select their features from the more robust dimensions of contrast. Acting jointly, Feature Economy, Marked Feature Avoidance and Robustness predict that languages tend to organize their sound systems in terms of a small number of highly-valued features, favoring unmarked feature combinations.

8. Phonological Enhancement

Enhancement is the name given to the reinforcement of weak acoustic contrasts by increasing the acoustic difference between their members (Stevens, Keyser & Kawasaki 1986, Stevens & Keyser 1989). Enhancement is feature-based as it typically affects natural classes of sounds rather than individual segments. Keyser & Stevens (2001) distinguish between Phonological Enhancement, in which reinforcement is achieved by introducing a redundant feature, and Phonetic Enhancement, in which reinforcement is achieved by introducing a supplementary articulation at the phonetic level. We shall be concerned with Phonological Enhancement here.

Phonological Enhancement typically involves the introduction of a marked feature value to reinforce an existing contrast between two classes of sounds. A familiar example is the assignment of the feature [+rounded] to back vowels. The introduction of this feature has the effect of lowering the second formant (F2) of back vowels, increasing their auditory distance from front vowels, which are characterized by a high F2. It will be recalled from Table 4 that [+rounded] is the marked value of the feature [±rounded]. In languages having this enhancement process, the marked value [+rounded] will be more frequent than the unmarked value [-rounded] in the class of back vowels, creating systematic violations of Marked Feature Avoidance.

The enhancement of posterior stops by the feature [+strident] is a further example of this type. The addition of [+strident] to a posterior stop increases its auditory distance from a nonstrident anterior stop such as /t/. In this case, the increase is not along a uniform auditory dimension, but along a different one. This is because /tʃ/ differs from /t/ not only in terms of its lower burst and transition frequencies, which depend on the feature value [+posterior], but also in terms of the presence of high-pitched, high-amplitude turbulence noise following the burst, which depends on [+strident].[24]

It is not always clear whether a given enhancing property is due to a feature operating at the phonological level or to a gesture introduced at the phonetic level. In the case of English /tʃ/, /dʒ/, however, the enhancement is clearly due to the feature [+strident], for three reasons. First, the stridency following the release of /tʃ/, /dʒ/ is not variable or gradient, but appears to be similar in duration, prominence and consistency to the distinctive stridency of affricates in languages in which they contrast minimally with *non*strident stops. Second, though this feature is redundant in the stops /tʃ/, /dʒ/ in English, it is distinctive in the fricatives /s/, /z/, which are minimally distinguished from /θ/, /ð/ by their stridency. The redundancy rule introducing [+strident] in the stops thus introduces a feature which is already distinctive in the system. Such "locally redundant features" (i.e., features redundant in some segments but distinctive in others) engage in Feature Economy effects just as fully distinctive features do (Clements 2003a), showing that they are phonologically present. Third, redundant values of [+strident] function in exactly the same way as distinctive values of [+strident] in English phonology. For example, both trigger an epenthetic vowel [ɨ] before the plural marker /-z/; compare nouns such as *matches* [...tʃɨz], with redundant [+strident], and *places* [...sɨz], with distinctive [+strident].

Examples of further common enhancement effects are shown in (31). All involve the introduction of a marked feature, in violation of Marked Feature Avoidance.[25]

(31) a. [+strident] enhances [+continuant] in coronal obstruents; thus the strident /s/, with its high-energy noise component, is more distinct from nonstrident stops like /t/ than is a nonstrident fricative like /θ/.

b. [+nasal] enhances [-continuant] in sonorant consonants; thus the nasal stop /n/, with its pronounced nasal resonance, is more distinct from continuants like /ɾ/ or /ɹ/ than is an oral noncontinuant like /l/.

c. [+posterior] enhances [coronal] in sonorant continuants; thus the palatal glide /j/, with its extra-high F2, is more distinct from noncoronals like /w/ than is a dental or alveolar continuant like /ɾ/ or /ð̞/.

d. [+labiodental] enhances [+continuant] in labial sounds; thus the labiodental fricative /f/, with its higher-amplitude fricative noise component, is more distinct from stops like /p/ than is a bilabial fricative like /ɸ/.

These examples are only illustrative, and others can be added.[26] All these enhancement effects can be expected to create reversals of the frequency patterns predicted by Marked Feature Avoidance.

That they do so is shown in Table 5, which illustrates all cases of Phonological Enhancement in consonants discussed up to now. The parenthesized numbers following a feature show the numbers of UPSID languages having that feature in the class of consonants described in the first column. For example, line 1 shows that 450 languages have [-strident] anterior coronal stops. The final, boldfaced lines in each set represent enhancement contexts. They show that in these contexts, marked values are *more* frequent than unmarked values. (Phonetic symbols illustrate typical realizations.)

in the class of:	*more frequent:*	*less frequent:*
a. anterior coronal stops	t [-strident] (450)	ts [+strident] (148)
posterior coronal stops	tʃ **[+strident]** (235)	c [-strident] (138)
b. coronal stops	t [-strident] (450)	tʃ [+strident] (291)
coronal fricatives	s **[+strident]** (397)	θ [-strident] (105)
c. vowels	a [-nasal] (451)	ã [+nasal] (102)
obstruents	t [-nasal] (451)	nt [+nasal] (57)
sonorant continuants	r [-nasal] (345)	nr [+nasal] (2)
sonorant noncontinuants	n **[+nasal]** (435)	l [-nasal] (368)
d. obstruents	t [-posterior] (450)	c [+posterior] (355)
sonorant noncontinuants	n [-posterior] (438)	ɲ [+posterior] (202)
sonorant continuants	j **[+posterior]** (384)	r [-posterior] (320)
e. labial stops	p [-labiodental] (446)	pf [+labiodental] (7)
labial sonorants	β [-labiodental] (34)	ʋ [+labiodental] (7)
labial fricatives	f **[+labiodental]** (199)	ɸ [-labiodental] (82)

Table 5. Frequency reversals resulting from Phonological Enhancement in consonants. Figures in parentheses show number of languages having each type of sound in the context shown at the left (source: UPSID). Enhancement contexts are shown in boldface.

The first two lines in this table show, for example, that while [-strident] coronal stops are present in more languages (450) than are [+strident] stops (148), the situation is reversed in the class of posterior stops, where [+strident] sounds hold the lead (235 vs. 138). The other boldface lines show similar frequency reversals.

The results for [±nasal] sounds in (c) are of particular interest. As noted earlier, [+nasal] is the marked value of the nasal/oral dimension by most criteria, yet nasal sonorants are more frequent across languages than oral sonorants. Closer study shows that this effect is due mainly to the predominance of nasals in the subclass of sonorant noncontinuants, which contains nasal stops and laterals. Within this subclass, as shown in (c), 435 languages have nasal or nasalized sounds while only 368 have [-nasal] sounds (laterals). Here it appears that [-continuant] is enhanced by [+nasal]. The explanation proposed in (31b) is that nasal stops like /m/ and /n/, with their pronounced nasal resonance, are more distinct from oral continuants like /ɾ/ or /ɹ/ than are oral noncontinuants like /l/.[27]

If this view is correct, it makes a strong prediction: nasal noncontinuants like /n/ should be much more frequent than oral noncontinuants like /l/ in languages having an oral continuant like /ɾ/. This prediction is borne out by an examination of the 320 UPSID languages which contain at least one R-sound (defined as any anterior coronal oral sonorant continuant) and one additional noncontinuant sonorant series, either nasal or lateral. In such "R-systems", as we might call them, the noncontinuant is almost invariably a nasal rather than a lateral, as shown below:

- 54 R-systems have an anterior coronal nasal N but no anterior coronal lateral L
- just 2 R-systems have an anterior coronal lateral L but no anterior coronal nasal N

This trend is highly significant (χ^2=20.446, p<.0001). Let us consider the two R-systems that do not conform to it. Waris (a Papuan language), has no plain nasal stops, but has prenasalized stops. In this system, enhancement of the noncontinuant /l/ with nasality would help distinguish it from the r-sound, but would tend to make it more similar to the prenasalized stop /nd/. Mixtec (a language of Mexico) has no coronal nasal stop but has a velar nasal and a series of nasalized vowels; here, too, coronal nasals would be competing with other nasal or nasalized sonorants.

While there is a similar trend in systems lacking R-sounds, this trend is much weaker. Of the 131 systems lacking R-sounds,

- 25 have an anterior coronal nasal N, but no anterior coronal lateral L
- 5 have an anterior coronal lateral L but no anterior coronal nasal N

This trend does not reach statistical significance (p>.05). Thus the preference for nasal sonorants is largely due to their overwhelming preponderance in R-systems having just one other anterior coronal sonorant series, as the enhancement-based account predicts.

We therefore find that a number of exceptions to the predictions of Marked Feature Avoidance can be explained in terms of Enhancement Theory. This result has an important consequence. While most earlier work on inventory structure (including Clements 2001) has interpreted frequency reversals such as shown in Table 5 as “markedness reversals” in which the marked and unmarked values of a feature are reversed in certain contexts, Enhancement Theory allows us to maintain that the marked value of a feature is the same in all contexts, even those in which it is most frequent. In this view, for example, the value [+nasal] is marked even in the class of sonorant noncontinuants where it is more frequent than [-nasal]. Independent markedness criteria support this view; thus [+nasal] is the typical (and perhaps unique) spreading value of [±nasal] regardless of the class of sounds in which it occurs. Had we maintained that [+nasal] is the unmarked value in sonorant continuants, we would have to maintain that the unmarked value spreads in just this special case.

There is one further implication of Enhancement Theory that is worth pointing out here. We have reviewed a number of cases in which marked feature values enhance contrasts in certain contexts. In other contexts, however, marked feature values have the effect of *reducing* contrasts, where they tend to be avoided. This fact helps to explain the avoidance of certain feature combinations which cannot be entirely explained in terms of Marked Feature Avoidance. For example, [+nasal] is widely disfavored in obstruents, even though obstruents are defined by the unmarked value [-sonorant]. Marked Feature Avoidance would predict instead that [+nasal] is avoided in the marked class of [+sonorant] sounds, where they are actually favored. The explanation here seems to be that [+nasal] reduces the auditory distinction between obstruents and sonorants when added to obstruents. While such *de-enhancement* effects have traditionally been dealt with as a further aspect of Markedness Theory, they seem more appropriately viewed as a complementary aspect of Enhancement Theory, in terms of which they receive a principled explanation.

To summarize this section, Phonological Enhancement is a principle by which weak acoustic contrasts are reinforced by redundant features. When the enhancing features are marked, the frequencies expected under Marked Feature Avoidance are typically reversed. In this way, Enhancement Theory accounts for a number of regular exceptions to the predictions of Markedness Theory.

9. Illustrations

We have discussed five feature-based principles that account for major trends in the structure of phonological inventories. One, Feature Bounding, is a bounding principle which sets the limits within which the others act. The other four are forces which interact with each other to produce systems which exhibit their effects to varying degrees. Exceptions to one tend strongly to reflect the operation of another.

Let us now see how these forces interact to distinguish likely from unlikely consonant systems. We consider small-inventory systems first. Under the principles discussed here, the same core set of relatively unmarked sounds should tend to be present in all systems, regardless of their size. This seems to be largely true.

(32) shows the ten commonest consonant types across the UPSID phoneme inventories.

(32)	P	T		K	
		S			
	M	N			
	W	L~R	J		H~ʔ

Taken as a possible small-inventory system, (32) is not implausible. It reflects Feature Economy at labial and coronal places of articulation. It makes use of unmarked values (except where Enhancement is involved, as in the case of the [+strident] S), and selects its contrasts from groups (a)-(d) of the Robustness Scale (26).

Small-inventory stems typically have most sounds of this set and sometimes add other members of the Basic Inventory (24), such as voiced stops. However, they do not add highly marked sounds, in conformity with Marked Feature Avoidance (18). Six small-inventory languages are shown in (33).

(33) Rotokas (Papuan)

p	t	k	
β		g	
	r		

Hawaiian (Austro-Tai)

p		k	ʔ
m	n		
w	l		h

Pirahã (Paezan)

p	t	k	ʔ
b		g	
	s		h

Roro (Austro-Tai)

p	t	k	ʔ
b			
m	n		
r	h		

Gadsup (Papuan)

p	t	k	
β	d		
m	n		
	j		ʔ

Maxakalí (Macro-Ge)

p	t	tʃ	k	ʔ
mb	nd	ndʒ	ŋg	
				h

Consider next larger-inventory systems. Such systems tend to conform to the predictions of Feature Economy, Marked Feature Avoidance, Robustness, and Phonological Enhancement. Rather than examining representative examples of larger-inventory systems conforming to these principles, however, since these are commonplace, let us look at a number of hypothetical systems *violating* these principles. Several are shown in (34)-(37).

(34) System A: violates Feature Economy

p	t	tʃ	k	
ɓ				
	t^h			
		dʒ		
			k’	
m				
	l			
		j		
			x	
				h

This system contains the high-frequency consonants /p t tʃ k/. However, it uses features uneconomically, as it has just one implosive, one aspirated stop, one voiced stop, one ejective, one fricative, one nasal, one liquid, one glide, and one glottal. Using 12 features to characterize 13 segments, it achieves an economy index of only 1.1.

(35) System B: violates Markedness

p	t	k	ʔ
b	d	g	
b^h	d^h	g^h	
v	z	ɣ	
m̥	n̥		
	l		
w	j		

This system, too, contains a core set of high-frequency consonants, and is relatively economical, but violates Marked Feature Avoidance (11). Thus, within UPSID languages,:

- voiced fricatives usually imply the corresponding voiceless fricatives
- voiceless nasals strictly imply voiced nasals[28]
- voiced aspirates strictly imply voiceless aspirates, both within UPSID and the data base of Indian languages collected by Pandey (2003).

The problem is that the gaps in system B correspond to unmarked rather than marked feature values.

(36) System C: violates Robustness

p	t	tʃ	k	ʔ
b	d	dʒ	g	
f	s	ʃ	x	h
v	z			

System C, like the others, contains a core set of high-frequency consonants. It is relatively economical and presents no significant violations of markedness. However, it violates the Robustness principle ("select higher-ranked features before lower-ranked features") as it has no sonorant consonants. Only two UPSID languages lack sonorant consonants and these are the small-inventory languages Pirahã and Maxakalí, shown in (33).

(37) System D: violates Enhancement

p	t	t̠	k
b	d	d̠	g
ɸ	θ	ç	x
β	ð		
w	ɾ l		h

This system satisfies previous criteria on most counts, but systematically fails to enhance weak contrasts. It chooses bilabial and nonsibilant fricatives instead of the preferred labiodentals and sibilants, incurring poor contrasts between e.g. /b/ and /β/, /β/ and /w/, /ɸ/ and /θ/. It selects nonsibilant posterior stops /t̠, d̠/ instead of the more distinctive sibilants /tʃ/, /dʒ/. It selects the oral sonorant /l/ instead of the nasal /n/, which provides a better contrast with /ɾ/.

To summarize, we find that the principles discussed earlier operate together to correctly describe small-inventory consonant systems and to exclude many imaginable, but unlikely larger-inventory systems. It appears that with nothing much more complicated than a ranked list of features indicating marked values, together with principles of economy and enhancement, we are in a good position to predict the preferred design features of phonological inventories to a reasonable first approximation.

10. Summary and Discussion

This paper has offered evidence that phonological inventories are structured by five feature-based principles. One, Feature Bounding, sets upper limit on the number of contrastive sounds that a language may have, while the others, Feature Economy, Marked Feature Avoidance, Robustness, and Phonological Enhancement, represent interacting forces that together define the set of preferred phonological systems within the limits set by Feature Bounding. The interaction of these principles accounts for the main design features of sound systems at the level where distinctive contrasts are taken into account. While these principles have been illustrated primarily with consonants here, they appear to hold for vowels as well; for example, the typical "symmetry" of vowel systems reflects Feature Economy. However, full discussion of vowels will require a separate study.

These results bear on the nature of the phonology/phonetics interface. Let us consider two theories of how phonology can be understood as constrained by phonetic factors: 1) a "direct access" theory, in which phonological generalizations make direct access to the near-infinite number of articulatory and acoustic parameter values provided by phonetic theory; 2) a "feature-mediated" theory, in which phonetics constrains phonology through the mediation of the phonetic definitions associated with a small set of distinctive features. This paper offers support for the second of these views: the major generalizations governing phonological inventories appear best captured in terms of principles stated over the features of which speech sounds are composed.

While this result confirms the central role of features in the organization of phonological inventories, it does not diminish the role of phonetics or of quantitative methods in understanding phonology. This is for at least three reasons. First, the predictions of these general principles must be fine-tuned by quantitative modeling in order to determine the relative weight of each and the precise nature of their interaction. This cannot involve a simple ranking, as no single principle (setting aside Feature Bounding) ever outranks all others. Second, these principles must be complemented by principles operating purely at the level of phonetic realization. These include:

- a theory of Gesture Economy, according to which sounds of a given class tend to have uniform gestural realizations (Maddieson 1995, Keating 2003); for example, anterior stops tend to be either dental or alveolar in the UPSID data base (Clements 2003a);
- a theory of Phonetic Enhancement, according to which weak feature contrasts may be enhanced by appropriate, subfeatural articulatory gestures; for example, palatoalveolar fricatives tend to be somewhat rounded in many languages to increase their auditory difference from anterior dental or alveolar fricatives (Keyser & Stevens 2001)

It is not clear at this point whether other proposed principles, such as a global measure of maximal dispersion, will be needed in addition to these. It is of course sound scientific practice to reduce explanatory principles to the necessary minimum.

Third, sound systems are what they are because speakers and hearers prefer sounds that are easily distinguished and not too hard to produce, a central insight of phonetic theory over many decades. Phonological inventories would be much different if linguistic expressions were realized in another medium. Indeed, studies of sign languages such as ASL confirm that the inventories of sign languages conform to quite different constraints. For example, as Sandler and Lillo-Martin (2001) have noted,

> Sign languages as well [as spoken languages] have constraints on the combination of elements .... For example, only one group of fingers may characterize the handshape within any sign. While either the finger group 5 (all fingers) or the group V (index plus middle finger) may occur in a sign, a sequence of the two shapes, *5-V is prohibited in the native signs of ASL and other sign languages.

This and many other constraints on sign language clearly are related to the specific nature of the medium and have no direct analogues in speech. However, it appears that sign languages, too, are characterized in terms of features (Brentari 1999), and one might ask whether the general principles discussed here generalize to basic design features of other linguistic media -- a topic which must be left for future work.

Why, ultimately, should phonological inventories be structured in terms of features rather than directly in terms of the finer-grained phonetic primes that define them? Here we may speculate that the answer may lie in quantal theory and the nature of early language acquisition, as reviewed in section 2. As was noted there, very young (including newborn) infants perceive speech sounds in terms of acoustic categories corresponding closely to the feature categories of adult languages, and are relatively insensitive to finer distinctions. These categories are typically determined by the natural boundaries that arise from the non-monotonic relationships between articulatory parameters and their acoustic effects.(as shown by many studies in quantal theory). It appears, then, that infants are biologically predisposed to perceive speech in terms of quantally-defined features. This implies that the ability to distinguish sounds across the more robust feature categories does not emerge during the early months of language acquisition, but is in place at the outset, constituting a perceptual "grid" within which speech information is processed. In the process of early language acquisition, this grid becomes *coarser* rather than more finely tuned, as categories that are not distinctive in L1 become merged. In short, speech is processed from the outset in a mode specifically adapted to feature categories which characterize the target language. If basic representational categories are fixed at a very early age, perhaps by the end of the first year (Juszcyk 1997; Peperkamp 2003), adult languages could be expected to preserve these categories, even though later phonetic training or prolonged exposure to another language may partly offset this effect.

What sort of consequences might we draw for the nature of phonological representations in the adult? The principles reviewed here suggest that as a minimum, such representations must contain marked distinctive features. As marked values represent a cost, there will be a tendency to minimize them (Marked Feature Avoidance). However, this cost is lessened to the extent that such feature values are supported by several sounds bearing them (Feature Economy). Fully redundant features will tend to be absent (Feature Economy, acquisitional merger). Speech sounds which are identical in feature terms are treated as equivalent (Feature Bounding). All else being equal, more robust contrasts are preferred to less robust contrasts (Robustness), and weak contrasts tend to be reinforced both phonologically and phonetically (Enhancement). Representational systems must be designed in such a way as to favor such characteristics.[29]

Acknowledgements. This paper has benefited from stimulating comments and suggestions received in the course of presentations in Paris, New York City, Ithaca, Copenhagen and Odense. I would particularly like to thank Andrea Calabrese, Morris Halle and Bert Vaux for their thoughtful comments on this paper at the Symposium on Phonological Theory, CUNY, New York, February 20-21, 2004.

Notes

[1] The notion "inventory" is understood here as an abstraction over set of distinctive segments (consonants and vowels) and subsegments (tones, autosegments) employed by a language in its phonological system, as defined by its active features.

[2] As Halle (1983) and Halle & Stevens (1991) have pointed out, the mapping between the articulatory feature definitions and their acoustic consequences is not always one-to-one. This is because, first, the acoustic realization of a feature may depend on other features of the sound in which it occurs. Thus, for example, the feature [stiff vocal cords] is realized as voicelessness in obstruents, but as high tone in vowels. Similarly, a given acoustic effect is not always associated with just one feature; e.g. voicing implements [slack vocal cords] in obstruents but [+sonorant] in vowels, regardless of their tone.

[3] Speech is not processed *only* in categorical terms. Speakers obviously also have knowledge of noncategorial aspects of speech sounds including intonation, voice quality, loudness, speech rate, and individual differences among speakers. Furthermore, fluent speakers of a language master fine details of phonetic realization which often differ considerably from one language to another. Categorical perception appears to be related to a phonological, as opposed to acoustic, mode of speech processing in which the listener processes the speech signal as speech rather than as noise. Recent physiological studies give evidence that both modes of processing are used in listening to speech. For example, studies of adults by Näätanen et al. (1997) have shown that phonemic representations distinct from those used in acoustical processing are based on a neural network predominantly located in the left temporal lobe. Electrophysiological studies of infants described by Dehaene-Lambertz & Pena (2001) suggest that categorical representations are not computed after acoustical representations, but in parallel with them.

[4] The actual number of sounds will be somewhat lower since in most feature systems, not all features combine freely. For example, in systems using [±high] and [±low], [+high] does not combine with [+low].

[5] Here and below, binary features are named by their marked values, where these are known. For reasons that will become evident, "posterior" is the marked value of the anterior/posterior feature, and so this feature is renamed [±posterior] here.

[6] Of course, caution must be exercised in taking articulatory labels found in the descriptive literature at face value. It is often hard to determine the exact value of sounds described as "dental", "alveolar", "palatal", and so forth, as such terms are often used impressionistically, or with different meanings from one writer to another.

7 Ladefoged & Bhaskakarao state that they are "unable to formulate phonetic criteria for deciding whether differences between sounds could be used phonologically within a language" (1983, 300).

8 This analysis presupposes a feature hierarchy in which [±sonorant] and the place features [labial], [coronal], and [dorsal] outrank [±voice], [±continuant], [±nasal] and [±lateral]. See further discussion in section 7.

9 It is assumed that the feature [coronal] is not usually required to distinguish coronal sounds from labial and dorsal sounds, for reasons given in Clements (2001). This feature is, however, phonologically active in English (Mohanan 1991, McCarthy & Taub 1992), and is assumed to be present as an active but nondistinctive feature in the phonological component.

10 This feature is probably [-obstruent] (Clements 2003a). Note that the devoicing of *k* did not lead to a merger with the original *k*, perhaps because it was largely confined to affixes.

11 The expected frequency for any cell is given by the formula $(T_R*T_C)/T_S$, where T_R and T_C are the row and column totals corresponding to the cell in question and T_S is the total sample size.

12 A strong advantage of a frequency-based approach is that it is defined on broad tendencies across languages while still allowing for language-particular variation. In some languages, further factors may increase the frequency of a marked sound, as in the case of the fricative [ð] in English, whose unexpectedly high frequency results from its occurrence in a small number of demonstratives (*the, this, then,* etc.) as well as from economy effects exerted by its voiceless counterpart [θ]. The high frequency of [ð] need not necessarily be taken as evidence that voicing is unmarked in English fricatives. However, the high frequency of a sound may facilitate its acquisition by language learners, and some research has suggested that markedness and frequency factors interact to account for variation in phonological acquisition (Stites et al. 2003).

13 Other evidence supporting the view that [coronal] is unmarked is summarized in Paradis & Prunet (1991). See Hume & Tserdanelis (2002) and Hume (2003), however, for evidence that non-coronal places of articulation show unmarked behavior in some languages. Yoneyama et al. (2003) report that [dorsal] is a more frequent place feature than [coronal] in Japanese, and that dorsal consonants tend to be acquired before coronal consonants, suggesting that [dorsal] may be less marked than [coronal] in that language..

14 Lombardi (1994) and Clements (2003a) argue that voiceless sonorants bear the feature [spread glottis]. If this analysis can be maintained in all languages, voicing need never be considered a distinctive feature of sonorants.

15 Some confusion has been created by the use of inconsistent transcription systems for variably voiced stops. As Dixon has observed, "A great deal of argument has gone on concerning whether *p, t, k* or *b, d, g* are most appropriate (at one time there was something in the nature of

a feud, triggered by this issue, between the 'voiceless symbol' Adelaide school and the 'voiced symbol' Sydney school" (Dixon 1980, 138). Practical orthographies sometimes employ one series of symbols and sometimes the other. For such reasons, it is dangerous to rely on transcription symbols – as does UPSID in one case -- in determining whether an Australian stop is voiced or voiceless.

[16] Such an analysis could also be expressed by underspecifying [±voiced] in these stops.on the assumption that nondistinctive values are phonetically underspecified in some languages (Keating 1988).

[17] However, not all exceptionless implicational statements are of the type discussed here, in which the implied term occurs in all languages. For example, the statement "dorsal fricatives imply dorsal stops" holds in all UPSID languages even though one UPSID language (Vanimo) lacks dorsal stops and fricatives. Here, in contrast to the statements in (13), the truth of the implication does not follow from the universal presence of the implied term. However, it does follow from the independent frequencies of dorsal stops (99.8%) and dorsal fricatives (33.6%) in UPSID, on the assumption that these sounds combine freely in languages. The expected frequency of systems having dorsal fricatives but lacking dorsal stops is a meager 0.07% (the product of .336 and 1-.998), which predicts that exactly .32 UPSID languages, or zero after rounding, should be of this type. Feature Economy further weighs against systems of this type, making them extremely improbable on statistical grounds. Implicational universals are not required in such cases to predict their quasi-absence.

[18] See section 8, however, for discussion of a significant class of majority patterns.

[19] This is perhaps the basic insight of Dispersion Theory, as developed in regard to vowel systems. In practice, however, Dispersion Theory has sought to define dispersion almost exclusively in terms of a two-dimensional (or sometimes three-dimensional) auditory space defined by formant values, and has not been successfully extended to the study of nonmodal vowels, or consonants. It is suggested here that the appropriate way of doing so is in terms of the principles of Robustness and Enhancement.

[20] The robustness of a contrast also varies according to the context in which it occurs. For example, place of articulation contrasts tend to be most robust in prevocalic position, and may be neutralized in contexts where their auditory cues are weak (Steriade 2001), unless they are enhanced by secondary cues. The Robustness Scale reflects preferred contrasts in their most favored contexts.

[21] As before, upper-case symbols stand for feature-defined classes of consonants rather than particular phonetic values. E.g. T = any voiceless coronal stop (whether dental, palatal, labialized, ejective, geminate, etc.). "Sonorant" = sonorant consonant. The feature [±posterior] is restricted, of course, to coronal sounds.

[22] UPSID does not provide information that would allow us to determine the frequency of contrasts based on the feature [±distributed], referring to the apical/nonapical distinction. This feature is often not recorded in primary descriptions, in part because relatively few languages make use of it for distinctive purposes in anterior consonants. See Calabrese (1994) for a proposed preference scale for vowel features.

[23] However, not all gaps involve segments that cumulate marked feature values. For example, most languages lack nasalized obstruents. In this case, though [+nasal] is a marked value, [-sonorant] is not. This and many other such gaps can be accounted for under Enhancement Theory, as discussed in section 8.

[24] More exactly, in [+posterior] obstruents the lowest spectral prominence is associated with F3 of neighboring vowels, while in [-posterior] obstruents this prominence is associated with F4 or F5 of neighboring vowels (Stevens 1989). As for [±strident], the high-frequency spectral energy of [+strident] sounds exceeds that of neighboring vowels, while the spectral energy of [-strident] sounds is lower than that of neighboring vowels at all frequencies (Stevens 1983).

[25] See further discussion and examples in Stevens, Keyser & Kawasaki (1986) and Stevens & Keyser (1989). In (31c), the use of [+posterior] to enhance [+continuant] sonorants suggests that [±posterior] may be the lower-ranked of these two features, in spite of its nearly equal frequency across languages. In (31d), [+labiodental] is used as an ad hoc feature to distinguish labiodental and bilabial sounds, on the view that [+strident] is restricted to coronal sounds. However, some linguists continue to use [+strident] for this purpose.

[26] No data could be compiled for the features [±ATR] and [±distributed], which are not consistently recorded in UPSID. However, it is well known that in African languages using [±ATR] as a distinctive feature, high vowels tend to be [+ATR] and low vowels to be [-ATR] (e.g. Archangeli & Pulleyblank 1994). These values reinforce acoustic distinctions, since [+ATR] and [+high] are realized with low F1 values and [-ATR] and [+low] with higher F1 values.

[27] The preference for a nasal series over a lateral series is also explained by feature economy: while [+nasal] can be applied to all major places of articulation, [+lateral] is largely restricted to coronal places of articulation.

[28] An isolated exception is Trumai, an Equatorial language of Brazil, which has two nasals *m* and *n*, of which the first is said to be typically voiceless (Monod-Becquelin 1975).

[29] In Clements (2001) I have suggested a minimalist approach to phonological representation along these lines.

References

Anderson, Jennifer, James L. Morgan, & Katherine S. White. 2003. "A statistical basis for speech sound discrimination," *Language and Speech* 46,155-182.

Archangeli, Diana & Douglas Pulleyblank 1994. *Grounded Phonology*. Cambridge, Ma.: MIT Press

Basbøll, Hans. 1985. Review of *Patterns of sound* (Ian Maddieson), *Phonology Yearbook* 2, 343-353.

Bell-Berti, Fredericka. 1993. "Understanding velic motor control: studies of segmental context." In M. K. Huffman & R.A. Krakow, eds., *Nasals, Nasalization, and the Velum*. San Diego: Academic Press, 63-85.

Best, Catherine T., G.W. McRoberts, & E. Goodell. 2001. "Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system," *Journal of the Acoustical Society of America*109.2, 775-794.

Best, Catherine T., McRoberts, G. W. & Sithole, N. M. 1988. "Examination of the perceptual reorganization for speech contrasts: Zulu click discrimination by English-speaking adults and infants," *Journal of Experimental Psychology: Human Perception and Performance* 14.3, 345-60.

Boersma, Paul. 1997. "Inventories in functional phonology," *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21, 59-90. (Reprinted as chapter 16 of *Functional Phonology*. The Hague: Holland Academic Graphics, 1998)

Bothorel, André. 1969-70. "Contribution à l'étude descriptive des latérales de l'albanais," *Travaux de l'Institut de Phonétique de Strasbourg* 2, 133-44.

Brentari, Diane. 1998. *A Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press.

Browman, Catherine P. & Louis M. Goldstein. 1989. "Articulatory Gestures as Phonological Units," *Phonology* 6.2. 201-251.

Browman, Catherine P. & Louis M. Goldstein. 1992. "Articulatory Phonology: an Overview," *Phonetica* 49, 155-180.

Browman, Catherine P. & Louis M. Goldstein. 2000. "Competing constraints on intergestural coordination and self-organization of phonological structures," *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5:25–34.

Calabrese, Andrea. 1994. "A constraint-based theory of phonological inventories." In W.U. Dressler, M. Prinzhorn & J.R. Rennison, eds. 1994. *Phonologica 1992: Proceedings of the 7th International Phonology Meeting*. Torino: Rosenberg & Sellier, 35-54.

Calabrese, Andrea. 1995. "A constraint-based theory of phonological markedness and simplification procedures," *Linguistic Inquiry* 26.3, 373-463.

Casali, Roderic F. 2003. " [ATR] value asymmetries and underlying vowel inventory structures in Niger-Congo and Nilo-Saharan, " *Linguisitc Typology* 7, 307-82.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Clements, G. N. 1999. "Affricates as noncontoured stops." In O. Fujimura, B. Joseph & B. Palek, eds., *Proceedings of LP '98: Item Order in Language and Speech*. Prague: The Karolinum Press, 271-299.

Clements, G. N. 2001. “Representational economy in constraint-based phonology.” In T. Alan Hall (ed.), *Distinctive Feature Theory*, 71-146. Berlin: Mouton de Gruyter.

Clements, G.N. 2003a. "Feature economy in sound systems," *Phonology* 20.3, 287-333.

Clements, G. N. 2003b. "Feature economy as a phonological universal." In M.J. Solé, D. Recasens & J. Romero, eds., *Proceedings of the 15th International Conference of Phonetic Sciences, Barcelona, 3-9 August 2003*, 371-374. Rundle Mall: Causal Productions Pty Ltd, 2003.

Clements, G. N. & Elizabeth Hume. 1995. “The internal structure of speech sounds.” In John Goldsmith, ed., *Handbook of Phonological Theory*. Oxford and Cambridge, MA.: Basil Blackwell, 245-306.

Clements, G. N. & Sylvester Osu (to appear) "Nasal harmony in Ikwere, a language with no phonemic nasal consonants," *Journal of African Languages and Linguistics 26.2*

Colarusso, John. 2004. “Georgian and other Caucasian languages,” unpublished ms., McMaster University.

Dehaene-Lambertz, G. & M. Pena. 2001. "Electrophysiological evidence for automatic phonetic processing in neonates," *Neuroreport* 12, 3155-3158.

Delgutte, Bertrand & Nelson Y.-S. Kiang. 1984. "Speech coding in the auditory nerve, IV. Sounds with consonant-like dynamic characteristics," *Journal of the Acoustic Society of America* 75, 897-907.

Dinnsen, D. A. 1992. "Variation in developing and fully developed phonetic inventories." In Ferguson, C.A., L. Menn & C. Stoel-Gammon, eds., *Phonological Development: Models, Research, Implications*. Timonium (MD): York Press, 191-210.

Dixon, R.M.W. 1977. *A Grammar of Yidiny*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 1980. *The Languages of Australia*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 1991. "Mbabaram." In R.M.W. Dixon & B. Blake, eds., *Handbook of Australian Languages*, vol. 4. Melbourne: Oxford University Press, 349-402.

Flemming, Edward. 2002. *Auditory Representations in Phonology*. New York & London: Routledge.

Greenberg, Joseph H. 1966. *Language Universals.* (2nd edition 2002.) Berlin: Mouton de Gruyter.

Groot, A.W. de. 1931. "Phonologie und Phonetik als Funktionswissenschaften," *Travaux du Cercle Linguistique de Prague* 4**.** 116-147.

Gurevich, Naomi. 2001. “A critique of markedness-based theories in phonology,” *Studies in the Linguistic Sciences* 31.2, 89-114.

Halle, Morris. 1983. “On distinctive features and their articulatory implementation,” *Natural Language and Linguistic Theory* 1, 91-105. (Reprinted in M. Halle, *From Memory to Speech and Back: Paper on Phonetics and Phonology 1954-2002*, Berlin: Mouton de Gruyter, 105-121)

Halle, Morris. 1992. "Phonological features." In William Bright (ed.), *Oxford International Encyclopedia of Linguistics*. N.Y.: Oxford University Press, 207-212

Halle, Morris. 2002. *From Memory to Speech and Back: Papers on Phonetics and Phonology 1954-2002*. Berlin: Mouton de Gruyter.

Halle, Morris & G. N. Clements. 1983. *Problem Book in Phonology*. Cambridge, Ma.: MIT Press and Bradford Books.

Halle, Morris & Kenneth N. Stevens. 1971. "A note on laryngeal features," *Quarterly Progress Report* 101, Research Laboratory of Electronics, MIT, Cambridge, MA, 198-213. (Reprinted in Halle 2002, 45-61.)

Halle, Morris & Kenneth N. Stevens. 1991. "Knowledge of language and the sounds of speech." In J. Sundberg, L. Nord, & R. Carlson, eds., *Music, Language, Speech, and Brain*. London: MacMillan Press, 1-19. (Reprinted in Halle 2002, 176-195.)

Harnad, Stevan (ed.) 1987. *Categorical Perception: the Groundwork of Cognition*. Cambridge: Cambridge University Press.

Hockett, Charles F. 1955. *A Manual of Phonology.* Memoire 11 of *IJAL* 21.4, Part 1. Baltimore: Waverley Press. Reprinted by the University of Chicago Press, 1974.

Hume, Elizabeth. 2003. “Language specific markedness: the case of place of articulation,” *Studies in Phonetics, Phonology and Morphology* 9.2, 295-310.

Hume, Elizabeth. & Georgios Tserdanelis. 2002. “Labial unmarkedness in Sri Lankan Portuguese Creole,” *Phonology* 19.3, 441-458.

Jakobson, Roman. 1968 [1941]. *Child Language, Aphasia and Phonological Universals*. (Translation of *Kindersprache, Aphasie und allgemeine Lautgesetze*, Uppsala, 1941) The Hague: Mouton & Co.

Jakobson, Roman & Morris Halle. 1956. *Fundamentals of Language*. The Hague: Mouton & Co.

Jakobson, Roman, Gunnar Fant & Morris Halle. 1952. *Preliminaries to Speech Analysis: the Distinctive Features and their Correlates*. (7th printing 1957) Cambridge, MA.: MIT Press.

Jusczyk, Peter W. 1997. *The Discovery of Spoken Language*. MIT Press, Cambridge, MA.

Jusczyk, Peter W., P.A. Luce & J. Charles-Luce. 1994. "Infants' sensitivity to phonotactic patterns in the native language," *Journal of Memory and Language* 33, 630-645.

Kean, Mary-Louise. 1980. The Theory of Markedness in Generative Grammar. (Doctoral dissertation, MIT, Cambridge, MA,1975) Distributed by the Indiana University Linguistics Club, Bloomington, Indiana.

Keating, Patricia. 1988. “Underspecification in Phonetics,” *Phonology* 5.2, 275-292.

Keating, Patricia. 2003. “Phonetic and other influences on voicing contrasts.” In *Proceedings of the 15th International Congress of Phonetic Sciences*. Rundle Mall: Causal Productions Pty Ltd., 375-378.

Kehrein, Wolfgang. 2002. *Phonological Representation and Phonetic Phasing: Affricates and Laryngeals*. Tübingen: Max Niemeyer.

Keyser, Samuel Jay & Kenneth N. Stevens. 2001. "Enhancement revisited." In Michael Kenstowicz, ed., *Ken Hale: a Life in Language*. Cambridge, MA: MIT Press, 271-291.

Kim, Hyunsoon. 2001. "A phonetically based account of phonological stop assibilation," *Phonology* 18.1, 81-108.

Kuhl, Patricia K., Karen A. Williams, Francisco Lacerda, Kenneth N. Stevens & Björn Lindblom. 1992. "Linguistic experience alters phonetic perception in infants by 6 months of age," *Science* 55, 606-608.

Ladefoged, Peter & Peri Bhaskakarao. 1983. "Non-quantal aspects of consonant production: a study of retroflex consonants," *Journal of Phonetics* 11, 291-302.

Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the World’s Languages.* Oxford: Blackwell Publishers.

Lang, Carrie E. & John J. Ohala. 1996. "Temporal cues for vowels and universals of vowel inventories," *Proceedings of ICSLP 96.*

Liberman, A.M., F.S. Cooper, D.P. Shankweiler, & M. Studdert-Kennedy. 1967. “Perception of the speech code,” *Psychological Review* 74, 431-461.

Liberman, A. M., K. S. Harris, H. S. Hoffman & B. C. Griffith. 1957. "The discrimination of speech sounds within and across phoneme boundaries," *Journal of Experimental Psychology* 54, 358-368.

Lieberman, Philip. 1970. Towards a unified linguistic theory, *Linguistic Inquiry*, **1,** 307-22.

Lindblom, Björn. 1986. "Phonetic universals in vowel systems." In *Experimental Phonology*, J. Ohala & J. J. Jaeger, eds., 13-44. Orlando: Academic Press.

Lindblom, Björn. 1992. "Phonological units as adaptive emergents of lexical development." In Charles A. Ferguson, Lise Menn & Carol Stoel-Gammon (eds.) *Phonological development: models, research, implications*. Timonium: York Press. 131-163.

Lindblom, Björn & Ian Maddieson. 1988. "Phonetic universals in consonant systems." In L.M. Hyman & C.N. Li, eds., *Language, Speech, and Mind*, Routledge, London, 62-78.

Lombardi, Linda. 1994. *Laryngeal Features and Laryngeal Neutralization*. (Ph.D. dissertation, University of Massachusetts, Amherst, Ma., 1991) New York: Garland.

McCarthy, J. & Alison Taub. 1992. Review of *The Special Status of Coronals* (Paradis & Prunet). *Phonology* 9, 363-370.

Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

Maddieson, Ian. 1991. "Testing the universality of phonological generalizations with a phonetically-specified segment database: results and limitations," *Phonetica* 48.2-4, 193–206.

Maddieson, Ian. 1995. "Gestural economy." In Kjell Elenius & Peter Branderud (eds.) *Proceedings of the 13th International Congress of Phonetic Sciences*, vol. 4, 574-577. Stockholm: Stockholm University.

Maddieson, Ian & Karen Precoda. 1989. "Updating UPSID," *UCLA Working Papers in Phonetics* 74. 104-111.

Martinet, André. 1955. *Économie des changements phonétiques*. Berne: A. Francke.

Martinet, André. 1968. "Phonetics and linguistic evolution." In Bertil Malmberg (ed.) *Manual of phonetics,* revised and extended edition, 464-487. Amsterdam: North-Holland Publishing Co. (Reprinted from the first edition edited by Louise Kaiser, 1957, 252-273.)

Maye, Jessica, Janet F. Werker, & LouAnn Gerken. 2002. "Infant sensitivity to distributional information can affect phonetic discrimination," *Cognitio*n 82, B101-B111.

Mohanan, K.P. 1991. "On the bases of radical underspecification", *Natural Language and Linguistic Theory* 9.2, 285-326.

Mohr, B. & W. S.-Y. Wang. 1968. "Perceptual distance and the specification of phonological features," *Phonetica* 18, 31-45.

Monod-Becquelin, Aurore. 1975. *La pratique linguistique des indiens trumai (Haut-Xingu, Mato Grosso, Brésil)*. Vol. 1. Paris: SELAF and CNRS.

Näätänen, Risto; Anne Lehtokoski, Mietta Lennes, Marie Cheour, et al. 1997. "Language-specific phoneme representations revealed by electric and magnetic brain responses," *Nature* 385, 432-4.

Pallier, Christophe, Laura Bosch & Núria Sebastián-Gallés. 1997. "A limit on behavioral plasticity in speech perception," *Cognitio*n, 64, B9-B17.

Pandey, Pramod. 2003. *Sounds and their patterns in the languages of India, Part II: Phonological sketches.* Ms., New Delhi: Jawaharlal Nehru University.

Paradis, Carole & Jean-François Prunet, eds. 1991. *Phonetics and Phonology*, vol. 2: *The Special Status of Coronals: Internal and External Evidence*. San Diego: Academic Press.

Peperkamp, Sharon. 2003. "Phonological acquisition: recent attainments and new challenges," *Language and Speech* 46, 87-113.

Pericliev, Vladimir & Raúl E. Valdés-Pérez. 2002. "Differentiating 451 languages in terms of their segment inventories," *Studia Linguistica* 56.1, 1-27.

Pierrehumbert, Janet. 2003. "Probabilistic phonology: discrimination and robustness." In R. Bod, J. Hay, & S. Jannedy, eds., *Probability Theory in Linguistics*. Cambridge, MA: MIT Press, 177-228.

Polka, L., & Werker, J. F. 1994. "Developmental changes in perception of non-native vowel contrasts," *Journal of Experimental Psychology: Human Perception and Performance,* 20, 421-435.

Repp, B.H. 1984. "Categorical perception: Issues, methods, and findings." In N. J. Lass, ed., *Speech and Language: Advances in Basic Research and Practice* (vol. 10). New York: Academic Press, 243-335.

Rice, Keren. 2000. "Featural markedness in phonology - Variation," *Glot International* 4.6, 3-6 and 4.7, 3-7.

Sagey, Elizabeth. 1990. *The Representation of Features in Nonlinear Phonology: the Articulator Node Hierarchy*. (1986 MIT PhD dissertation) N.Y.: Garland.

Sandler, Wendy & Diane Lillo-Martin. 2001. "Natural sign languages." In M. Aronoff & J. Rees-Miller (eds.), *Handbook of Linguistics*, 533-562.

Simpson, Adrian P. 1999. Fundamental problems in comparative phonetics and phonology: does UPSID help to solve them? In *Proceedings of the 14th International Congress of Phonetic Sciences*, vol. 1. Berkeley: University of California. 349-352.

Steriade, Donca. 2001. "Directional asymmetries in place assimilation: a perceptual account." In E. Hume & K. Johnson (eds.), *The Role of Speech Perception in Phonology*. New York: Academic Press, 219-250.

Stevens, Kenneth N. 1972. "The quantal nature of speech: evidence from articulatory-acoustic data." In E. E. David, Jr. & P. B. Denes, eds., *Human Communication: a Unified View.* New York: McGraw-Hill, 51-66.

Stevens, Kenneth N. 1983 "Design features of speech sound systems." In *The Production of Speech*, Peter F. MacNeilage (ed.), 247-261. New York: Springer-Verlag.

Stevens, Kenneth N. 1989. "On the quantal nature of speech," *Journal of Phonetics* 17**.** 3-46.

Stevens, Kenneth N. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Stevens, Kenneth N. & Samuel Jay Keyser. 1989. "Primary features and their enhancement in consonants," *Language* 65: 81-106.

Stevens, Kenneth N., Samuel Jay Keyser & H. Kawasaki. 1986. "Toward a phonetic and phonological theory of redundant features." In J. Perkell and D. Klatt, eds., *Symposium on Invariance and Variability of Speech Processes.* Hillsdale: Lawrence Erlbaum, 432-469.

Stites, Jessica, Katherine Demuth, & Cecilia Kirk. 2003. "Markedness vs. frequency effects in coda acquisition." To appear in *Proceedings of the Boston University Conference on Language Development,* 28.

Trehub, S.E. 1976. "The discrimination of foreign speech contrasts by infants and adults," *Child Development*, 47:466-472.

Triulzi, A., A.A. Dafallah & M.L. Bender. 1976. "Berta." In M.L. Bender (ed.), *The Non-Semitic Languages of Ethiopia*. East Lansing, MI: African Studies Center, Michigan State University, 513-32.

Trubetzkoy, Nicolai S. 1969. *Principles of Phonology.* (English translation of *Grundzüge der Phonologie,* Göttingen: Vandenhoeck & Ruprecht, 1939.) Berkeley and Los Angeles: University of California.

Werker, Janet F., J. H. V Gilbert, K. Humphrey, and R. C. Tees. 1981. "Developmental aspects of cross-language speech perception," *Child Development*, 52:349-355.

Werker, Janet F. & R. C. Tees. 1984. "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life," *Infant Behaviour and Development*, 7:49-63.

Westbury, John R. 1983. "Enlargement of the supraglottal cavity and its relation to stop consonant voicing," *Journal of the Acoustical Society of America* 73**,** 1322-1336.

Williamson, Kay. 1973. "More on nasals and nasalization in Kwa," *Studies in African Linguistics* 4: 115-138.

Wonderly, William L. 1946. "Phonemic acculturation in Zoque," *International Journal of American Linguistics* 12, 92-95.

Wonderly, William L. 1951-2. "Zoque (parts I-V)," *International Journal of American Linguistics* 17 (1951), 1-9, 105-123, 137-162, 235-251; *International Journal of American Linguistics* 18 (1952), 35-48.

Yallop, Colin. 1982. *Australian Aboriginal Languages*. London: André Deutsch.

Yoneyama, K., M.E. Beckman & J. Edwards. 2003. "Phoneme frequencies and acquisition of lingual stops in Japanese." Unpublished ms., Ohio State University.