University of Massachusetts Occasional Papers in Linguistics

Volume 19 *University of Massachusetts Occasional Papers in Linguistics -- Volume 16 -- Phonological Representations*
Article 4

1993

Size, Structure, and Markedness in Phonological Inventories

John Kingston
University of Massachusetts, Amherst

Recommended Citation

Kingston, John (1993) "Size, Structure, and Markedness in Phonological Inventories," *University of Massachusetts Occasional Papers in Linguistics*: Vol. 19 , Article 4.
Available at: https://scholarworks.umass.edu/umop/vol19/iss1/4

Size, Structure, and Markedness in Phonological Inventories*

John Kingston

Linguistics Department
University of Massachusetts, Amherst

0 Abstract

This paper employs statistical analyses of various co-occurrence dependencies between consonants differing in place and manner of articulation to argue that large inventories tend not differ from smaller ones merely in having more individual segments but rather in having entire blocks of sounds that are absent in smaller inventories. The blocks of sounds found in larger inventories increase the number of contrasts by combining articulations in such a way as to minimize articulatory cost while maximizing perceptual distance between contrasting sounds. The statistical analyses thus support Lindblom's (1986, Lindblom & Maddieson, 1988; also Ohala, 1979) theory of adaptive dispersion as a principle governing inventory size and structure. Furthermore, markedness as a property of inventories is redefined such that the more marked inventories are those with incomplete blocks of sounds rather than simply those with more sounds.

1 Introduction

The usual view of how a large inventory of speech sounds differs from a small one is that the larger one has added some number of marked sounds to a basic inventory of unmarked sounds, which makes up most or all of the small inventory (for a recent proposal along these lines see Lindblom & Maddieson, 1988).[1] Inventories of different sizes should furthermore differ incrementally in markedness from one another. In the first steps beyond the basic inventory of unmarked sounds, the additional sounds in slightly larger inventories should differ minimally in markedness from the basic sounds. But in yet larger inventories, sounds even greater markedness should appear, that are absent (or rare) in inventories of intermediate size (and markedness). On the other hand, markedness does not predict just what marked sounds will appear in an inventory of a certain size but not in a slightly smaller one. This paper presents evidence that both predictions of markedness theory are incorrect (at least in part).

First, larger inventories turn out to differ from smaller ones in having whole blocks of sounds that are absent from the smaller inventories. Discontinuous differences in inventory size are suggested by the significantly greater frequency of languages with all or none of the sounds in the block and the significantly lesser frequency of languages with only some of the sounds in the block. Thus, markedness does not increase incrementally but instead in jumps. The larger inventory resulting from adding an entire block of sounds tends to employ the new contrast(s) to the fullest extent possible (hence they will be called "contrast blocks" in the remainder of this discussion). For example, in the large inventory of English obstruents [voice] contrasts not only in the stops and affricates, /p:b/, /t:d/, /tʃ:dʒ/, /k:g/, but also in the fricatives /f:v/, /θ:ð/, /s:z/, /ʃ:ʒ/, even though voiced fricatives are rather rare, especially when also sibilant. Since inventories with or without complete contrast blocks are statistically more common, they may be closer to an equilibrium state than inventories with just a some of the sounds in a contrast block. And their greater equilibrium may actually reduce the markedness of the inventory as a whole compared to that of inventories with only partial contrast blocks, even though the inventories with complete contrast blocks are larger. Thus markedness of entire inventory may not be a monotonic function of its size.

Second, it is possible to predict with a certain degree of precision just what sounds will be found in an inventory of a certain size that will be absent from a slightly smaller one, by appealing to perceptually motivated articulatory covariation.

The evidence presented in support of both these claims comes from various statistical assessments of the distributions of place and manner contrasts among obstruents in the inventories collected in the UCLA Segment Inventory Database (UPSID; Maddieson, 1984). This paper shows furthermore that explanations for how larger inventories differ from smaller ones are to be found in speakers' deliberate combining of articulations whose acoustic effects interact perceptually in desirable ways.

Section 2 presents evidence showing that the requisite degree of articulatory control to produce such perceptually desirable combinations is available to and used by speakers. That section also explores differences between explanations based on speakers' attempts to influence perception and those based entirely on how articulations are organized and accomplished. Section 3 tests perceptual explanations against entirely articulatory ones in terms of how well each predicts the places at which languages will have stops for languages contrasting different numbers of places for stops. The hypothesis that speakers combine articulations to influence perception is also explored in an examination of the behavior of affricates (section 4) and fricatives (section 5) with respect to the number and kind of place contrasts in obstruents. Section 4 examines the use of a palato-alveolar affricate as an additional place of articulation for stops and its co-occurrence possibilities with other non-anterior non-continuants, as well as the dependencies between palato-alveolar affricates and other sibilant obstruents. Section 5 looks generally at dependencies among places in fricatives.

2 Perceptual explanations for articulatory covariation

2.1 Introduction

Many articulations covary in the production of any minimal contrast between speech sounds. And not only do articulations covary within single distinctive feature values, but entire sets of articulations also covary, in any redundancy between distinctive feature values. Most explanations of covariation, whether within or between distinctive feature values, attribute it to a mechanical yoking of the articulations to one another. However, the evidence frequently suggests instead that speakers may deliberately covary articulations to produce desirable perceptual interactions among their acoustic effects (Kluender et al., 1988, Diehl & Kluender, 1989; Diehl & Kingston, 1991; Kingston, 1991, 1992; Kingston & Diehl, submitted).

2.2 Covariation within a distinctive feature

To get an idea of how the perceptual and articulatory explanations differ, consider the covariation of soft palate height and rate of vocal fold vibration with tongue height in vowels: in vowels with higher tongue positions, the soft palate is also typically higher and the vocal folds vibrate faster. The standard articulatory accounts of these covariations treat them as mere mechanical side effects of raising or lowering the tongue. The soft palate can stay high in high vowels because the high tongue position removes any downward pull by the tongue on the soft palate via the palatoglossus (Moll, 1962; Moll & Shriner, 1967). On the other hand, raising the tongue may pull upward (and perhaps forward) on the hyoid bone which in turn pulls the larynx up (and forward); raising the larynx increases the vertical tension applied to the vocal folds and thus their rate of vibration (Honda, 1987; Honda & Fujimura, 1991; Ohala & Eukel, 1987).

Both of these accounts are seriously undermined by evidence that these palatal and laryngeal articulations are deliberately rather than passively produced. First of all, the levator palatini is more active in high than low vowels (Lubker, 1969). Since levator palatini contraction is the principal mechanism for raising the soft palate, it appears that soft palate is deliberately kept up in high vowels rather than staying high because no tongue lowering occurs to pull it downward via the palatoglossus. Similarly, the cricothyroid muscle is more active in high than low vowels (Vilkman, Aaltonen, Raimo, Arajärvu & Oksanen, 1989; Dyhr, 1991). Cricothyroid contraction is the principal mechanism for increasing the rate of vocal fold vibration (Ohala, 1970; Collier, 1974), by tilting the thyroid cartilage down and forward and thus stretching the vocal folds. Therefore, it appears that the folds' rate of vibration is deliberately rather than mechanically varied with tongue height.

But why should speakers deliberately raise the soft palate and increase the rate of vocal fold vibration in high vowels? The answer is that an elevated soft palate and higher rate of vocal fold vibration have acoustic effects which interact perceptually with those of the high tongue position so as to enhance the difference between high and low vowels..

First of all, the interactions among the acoustic effects of soft palate and tongue height suggest that the soft palate is actively kept high in high vowels so that nasalization does not distort the vowels' perceived highness and is allowed to lower (or to be pulled down mechanically) in low vowels because modest nasalization actually enhances the perceived lowness of the vowel. Besides a high F_1, low vowels are also characterized by a broader distribution of energy (a greater bandwidth) in the lower frequency range of their spectra. In high vowels, F_1 is instead low and low frequency energy is concentrated in a narrow band. The modest nasalization that comes from the slightly lower soft palate in low vowels, will raise F_1 and introduce a nasal pole-zero pair below F_1, while in high vowels, it raise F_1 and introduce the nasal pole-zero pair above the vowel's F_1 (Stevens, Fant & Hawkins, 1987). Thus moderate nasalization will enhance the acoustic effects of the low tongue position in low vowels, but distort, in the direction of a lower vowel, the effects of the high tongue position in high vowels.[2]

Second, a higher rate of vocal fold vibration will bring F_0 closer to the low F_1 of a high vowel. The close proximity of F_0 to low F_1 creates a low, narrow band of energy which contrasts with the higher and broader band of energy that results from the lower F_0 and higher F_1 of lower vowels. And Syrdal and Gopal (1986) have shown that the separation between F_1 and F_0 (calculated in Bark to reflect the nonlinear transformation of the acoustic signal in the peripheral auditory system) reliably distinguishes the high from non-high vowels in Peterson and Barney's (1952) sample of the vowels 76 men, women, and children.

For both soft palate height and rate of vocal fold vibration, their observed covariation with tongue height can be plausibly argued to arise because the speaker deliberately combines acoustic effects which will enhance the vowel height contrast. The acoustic effects of covarying articulations may even integrate into a single perceptual property: here, the effects of covarying soft palate height and vocal fold vibration with tongue height apparently produces a contrast between vowels with a broad, high band of energy at the bottom of the spectrum (low vowels) vs. vowels with a narrow, low band of energy (high vowels). (Diehl and Kluender (1989) and Kingston (1991) present further evidence regarding the nature of the perceptual interactions caused by covarying rate of vocal fold vibration and soft palate height with tongue height.)

2.3 Defining integration

The term "integration" will be used here to refer only to those cases where the acoustic effects of independent articulations merge into a single perceptual property and not to other kinds of perceptual interactions between the acoustic effects. In particular, if the value of one acoustic effect influences the judgment of another, yet the two effects are perceived separately, then their interaction is not integrative.

For example, if the F_2 of an adjacent vowel is low enough it will lead a listener to judge a relatively low frequency sibilant to be an [s] rather than an [ʃ] (Mann & Repp, 1980; Whalen, 1981, 1990, 1992; Nearey, 1990), presumably because the listener expects the sibilant's noise to be lowered in frequency by coarticulation with the adjacent vowel. Since the vowel's F_2 and the height of fricative noise are still perceived as independent properties (one for the backness of the vowel and the other for the anteriority of the fricative), this interaction is not integration in the sense this term is used here.

Other kinds of perceptual interactions, particularly those which have been called "trading relations" (Repp, 1982), may be integrative. For example, a sufficiently high F_1 at the onset of the following vowel may lead a listener to judge a stop with a relatively short VOT as voiceless aspirated because both effects lead diminish the percept of low frequency energy near the stop release. That is, VOT and F_1 onset frequency integrate into a percept of the proximity of low frequency energy to the stop release (Diehl & Kingston, 1991; Kingston & Diehl, submitted). Other acoustic effects trade without appearing to integrate, however. For example, Recasens (1983) has shown that rising F_2 and falling F_3 transitions combine with high frequency nasal formants and a high frequency nasal zero in the nasal murmur to convey a post-vocalic [ŋ], while rising F_2 and rising F_3 transitions combine with low frequency nasal formants and a low frequency nasal zero in the nasal murmur to convey a post-vocalic [n]. Because both acoustic effects are apparently necessary to distinguish [ŋ] from [n], Recasens describes the interaction of the formant transitions and the spectrum of the nasal murmur as integrative. But because these acoustic effects don't both contribute to any obvious single perceptual property, they would not be described as integrative here. (For other examples of perceptual interactions that have been described as integrative, but which don't apparently yield a single perceptual property, see Dorman, Studdert-Kennedy and Raphael (1977) and Repp, Liberman, Eccardt and Pesetsky (1978).)

2.4 Covariation between distinctive features (redundancy rules)

At least some instances of articulatory covariation of the sort expressed by redundancy rules may also be perceptually rather than articulatorily motivated. Consider the extraordinarily frequent covariation/-redundancy of lip rounding with tongue backness: well over 90% of the vowels in UPSID are either front, unrounded or back, rounded (see also Diehl & Kluender, 1989, for extensive discussion). Constrictions at the back of the palate will lower F_2, while those at the front will raise this formant. Lip rounding aids retracting of the tongue body in lowering F_2 in back vowels in two ways: by lengthening the vocal tract and constricting the mouth opening (both of these articulatory settings actually lower all the resonances of the vocal tract).[3] Lip spreading, on the other hand, aids advancing the tongue body in raising F_2 by shortening the vocal tract and flaring the mouth opening. The lingual and labial articulations in vowels thus regularly covary in this way because their acoustic effects are similar. As there is no mechanical connection whereby pulling the tongue back in the mouth rounds the lip, an articulatory explanation of this covariation can be safely rejected in this case (but see the discussion of global articulatory goals below).

The covariation of soft palate height and rate of vocal fold vibration with tongue height in vowels produces combinations of acoustic effects that enhance the height contrast. Other combinations of articulations appear instead to diminish the differences between contrasting sounds and so are avoided. The diminution may arise because the acoustic effects of the covarying articulations are so thoroughly integrated by listeners that they fail to adjust for one articulation' effect on another (see Kingston, 1992, for more extensive discussion).

One such case is the redundancy between [voice] and [continuant] in many languages, where [+continuant] obstruents are [-voice]. First, UPSID shows that [voice] contrasts are markedly less common in fricatives than stops, at least for fricative inventories of three or fewer. Second, [voice] contrasts are even rarer in sibilant than non-sibilant fricatives. Finally, voiced non-sibilant fricatives occur frequently without their voiceless counterparts (apparently through lenition of the corresponding voiced stops, Maddieson, 1984), while voiced sibilants occur only very rarely without their voiceless counterparts. These distributional facts suggest that a contrast for the laryngeal feature [voice] depends on the manner features [sonorant] and [continuant] and the place features [coronal] and [strident]. In [-sonorant, +continuant] sounds, [+voice] implies either [-coronal] or if [+coronal], then [-strident], while [+coronal, +strident] (= sibilant) implies [-voice].

Voiced sibilants are probably rare because adducting the glottis to get the vocal folds to vibrate reduces transglottal air flow enough that air flow through the constriction downstream is too low to produce sibilants' characteristic high intensity noise. The combination of [+voice, +continuant, +strident, coronal] yields an array of acoustic properties that overlaps too much with the array produced by [+voice, +continuant, -strident, coronal], making the collapse of the [strident] contrast probable. The contrast will only be maintained so long as the threshold for deciding that a fricative is [+strident] is lowered to accommodate the relative weakness of the noise in the production of voiced sibilants. But in order to lower the [+strident] threshold appropriately, the listener has to separate perceptually the acoustic effects of the articulations specified by [+voice, coronal] from those of [+strident], and use these separated percepts as criteria for assessing the acoustics of the [+strident] articulations. If [+voice, coronal] is instead integrated with [+strident], the [+strident] threshold will remain high, and the contrast may collapse in favor of the [-strident] sound.[4] /z/ should thus become /ð/, but since /ð/ itself is relatively rare, it appears that /z/ more often becomes /r/, which like /ð/ is [-strident].

Integration thus produces results similar to the kind of misperception (Ohala, 1981) that causes many sound changes. Misperception occurs when the listener fails to detect the coarticulatory source of an acoustic distortion of a segment and interprets it instead as an intended property of that segment. Perceptually separating the acoustics of [+voice, coronal] from [+strident] in voiced sibilants has the same effect as detecting the coarticulatory source of an acoustic distortion since it allows the threshold for judging stridency to be adjusted appropriately.

2.5 Motor equivalence and articulatory phonology's account of covariation

Besides the kind of mechanical yoking that was supposed to explain the covariations in soft palate height or rate of vocal fold vibration with tongue height, articulatory accounts of covariation have more recently recognized that some articulations covary because they all contribute to achieving the same articulatory goal, a property called "motor equivalence" or "equi-finality" (MacNeilage, 1970; Fowler, Rubin, Remez & Turvey, 1980). Two articulations are said to be "motor equivalent" if both contribute to and either may serve to accomplish a particular articulatory goal. For example, achieving a particular degree of constriction in articulating a vowel can be accomplished either by raising the tongue or by raising the jaw (Ladefoged, DeClerk, Lindau & Papcun, 1972). The motor equivalence of tongue and jaw raising was demonstrated most clearly by experiments where bite blocks were inserted between the teeth to prevent jaw raising in a close vowel (Lindblom, Lubker & Gay, 1979; Fowler & Turvey, 1980). In all such studies, speakers immediately compensated by raising the tongue more to produce the same degree of constriction at the same location on the palate and thereby achieved a vowel quality that was virtually indistinguishable from that produced when the jaw is free to aid tongue raising in achieving a close constriction.[5]

Motor equivalence among articulations is one of the two key properties of gestures' phonetic implementation in Browman and Goldstein's articulatory phonology (1986, 1988, 1990, in press).[6] The gestures themselves are the units of phonological representation in this theory and they encode articulatory goals via tract variables specifying the location and degree of constriction in the vocal tract. What articulations actually combine together to achieve these goals is not, however, directly specified in the tract variables; instead from moment to moment the task dynamics chooses the articulators and determines how far they will move based on all current gestural demands on them.

Leaving the selection of the actual articulations which achieve the goal specified by the tract variables entirely to the task dynamics insulates the tract variables from articulatory covariation, and prevents that variation from influencing the phonological representations. Thus in the eyes of the tract variables and the gestures composed of them, all patterns of articulatory covariation which achieve the goal they specify are equal. Therefore, the only way that one covariation pattern could be selected over others would be because that pattern conforms to how the covarying articulations mechanically influence and constrain each other. But since some instances of covariation do not arise from such mechanical effects, the task dynamics will need the help of additional constraints to select the more frequent covariation patterns over the less frequent.

These additional constraints come in a variety of forms. First, articulatory goals could be global as well as local changes in the vocal tract's configuration. That speakers may have global goals can be inferred, for example, from the fact that larynx lowers in rounded vowels to add additional length to a vocal tract already lengthened by lip protrusion (Lindblom & Sundberg, 1971). Increasing the length of the vocal tract lowers the frequency of all its resonances. Even more telling evidence that larynx lowering and lip protrusion are coordinated articulations comes from Riordan's (1977) demonstration that speakers lower the larynx more in rounded vowels when they are prevented from protruding the lips.[7] Second, as already noted, the possible perceptual interactions among the acoustic effects of covarying articulations are another, equally rich source of constraints on that covariation. Third, the exclusively articulatory constraints embodied in the task variables and dynamics of articulatory phonology are unfortunately still not specified precisely enough to distinguish all the places of articulation whose representation is the focus of this paper. That theory's predictions regarding what patterns of articulatory covariation should be observed cannot therefore be readily compared with the predictions stemming from perceptual interactions.

2.6 Feature geometry's account of covariation

Feature geometric representations, on the other hand, are sufficiently precise to make the necessary distinctions of place, and to predict the relative frequency of particular places. Feature geometry explains the ubiquity of stops at the three major places of articulation: bilabial, dental or alveolar, and velar, by specifying three articulator nodes: [labial], [coronal], and [dorsal]. Its representations also predict that additional places of articulation will be produced by means of finer or more specialized uses of these articulators, which are represented by the features subordinate to the articulator nodes. Feature geometry thus encodes the greater markedness of these additional and less frequent places directly in the hierarchical arrangement of articulator nodes and their subordinate features.

Although the most compelling support for feature geometry's hierarchical representations of the internal relations among distinctive features proposals comes from the grouping of sounds into natural classes by rules or co-occurrence restrictions (McCarthy 1988), a comprehensive model of sounds' phonological behavior should also account for the structure of inventories. When examined against the distributions of place contrasts among stops and fricatives in UPSID, the predictions of feature geometry prove unable to account comprehensively for the facts, and must be supplemented by the predictions of the hypothesis that perceptual interactions shape articulatory covariation.

The predictions of a theory based on perceptually motivated articulatory covariation regarding the relative markedness of inventories also differ from those of feature geometry. Feature geometry's representation of markedness, predicts a smooth and continuous rather than lumpy distribution of inventory sizes, and that the relative markedness of inventories will be incremental and monotonically increasing with inventory size.

3 Increasing the number of places of articulation for stops

3.1 Specializations of lingual articulations and combining labial and lingual articulations

All but 5 of the 317 languages in UPSID have stops at the three major places of articulation, bilabial, dental or alveolar,[8] and velar,[9] which accords with feature geometry's recognition of three major articulators: the lips, the tongue tip-blade, and the tongue dorsum, encoded as the articulator nodes [labial], [coronal], and [dorsal].

The remainder of this section examines the distribution of stops across places of articulation beyond the major three places. Place is discussed in the terms of first-generation feature geometric proposals such as Clements (1985), Sagey (1986), or McCarthy (1988), i.e. in terms of articulator nodes, which may dominate subordinate features (later generations of feature geometry such as those presented in Clements (1989) or Hume (1992) represent specialized places somewhat differently).

In languages with more than the three major places for stops, some of the additional places are specialized uses of the major articulators. Two patterns are observed: (1) contrasts between palatal or uvular and velar stops are specialized tongue body articulations, which are representable with the features [back] and [high], respectively, under the [dorsal] node and (2) contrasts between dental, retroflex, or palato-alveolar stops and alveolar stops are specialized tongue tip-blade articulations, which are representable with the features [distributed] or [anterior] under the [coronal] node.

The other means of producing additional places is to combine two articulators; the most common example of such double articulations use the [labial] and [dorsal] articulators. Other combinations where a complete closure is made with two articulators, such as labio- or dorso-coronals, are vanishingly rare, and they will be ignored here.[10] Doubly-articulated stops are included in this discussion under two definitions of what counts as doubly-articulated: (1) where each articulation achieves complete closure, as in /k͡p/, (table 1) and (2) where only one does, while the other remains more open, as in /kʷ/ -- the latter are usually treated as stops with secondary articulations (table 2).

Both tables 1 and 2 show that a specialized [dorsal] articulation, i.e. either a palatal or a uvular, occurs more than twice as often as a specialized [coronal] articulation in languages with just one additional place beyond the major three. In languages with two additional places, two

Table 1 Lgs. with 4-6 places: c palatal, q uvular, t̪ dental, t̄ palato-alveolar, ʈ retroflex, kp labio-velar. Art(iculator), B(ody), D(ou)bl(e), O(ther), Spec(ialization), T(ip-Blade).

Places:	Pal	Uvu	Dnt	Pal-Alv	Ret	Lab-Vel		
Spec/Dbl:	Body		Tip-Blade			Lips and Body		
Art:	[dorsal]		[coronal]			[labial] and [dorsal]		
	[-back]	[-high]	[+ant] [+dist]	[-ant] [+dist]	[-ant] [-dist]			
Places:	c	q	t̪	t̄	ʈ	kp		
Places								Lgs
4	34	30	7	4	18	12		
Art		64			29	12		106
5							Lgs	
B&T	5	7	5	0	7	0	12	
2B/2T	6	6	3	2	5	0	11	
O/Dbl	4	0	3	0	1	8	8	
Art		28			26	8		31
6*	6	0	6	2	4	0		6
Art		6			12	0		
Spec	51	43	21	8	34	20		
Art		94			63	20		
Total Lgs								143

specializations of either the [dorsal] or [coronal] articulators are about as common as one specialization of each of them. All languages with three additional places have one specialized [dorsal] articulation; otherwise, two distinct patterns emerge, depending on what counts as a doubly-articulated stop. When both constrictions must achieve complete closure (table 1), there are no languages with three additional places where any of them is doubly articulated. Furthermore, the specialized [dorsal] articulation is always a palatal, not a uvular, and thus the other two are specialized

Table 2 Lgs. with 4-6 places with common secondary articulations as separate places: k^w, q^w labialized velar, uvular; t^j, k^j palatalized alveolar, velar.

Spec/Dbl:		Body		Tip-Blade			Body or Tip/Blade and Lips or Body/Tip-Blade*				
Art:		[dorsal]		[coronal]			[dorsal] or [coronal] and [labial] or [dorsal]*				
		c	q	t̲	t̄	t	kp	k^w, q^w	t^j	k^j	
Places											Lgs
4		33	19	5	4	15	12	20	6	0	
	Lab/Pal							32		6	
	Spec/Dbl		52			24				38	114
5											
	B&T	5	5	4	0	6	0	0	0	0	10
	2B/2T	4	4	3	2	5	0	0	0	0	9
	D	5	2	5	0	4	8	7	5	6	14
	Ttl	14	11	12	2	15	8	7	5	6	
	Lab/Pal							15		11	
	Spec/Dbl		25			29				26	33
6		6	11	7	2	4	0	11	0	0	
	Lab/Pal							11		0	
	Spec/Dbl		17			13				11	
											17
	Ttl	53	41	24	8	34	20	38	11	12	
	Lab/Pal							58		23	
	Spec/Dbl		94			66				81	
	Ttl										164

* This assumes that the tongue body = [dorsal] articulator is used for palatalization; if it is tongue tip-blade = [coronal] articulator instead, then [dorsal] should be replaced by [coronal] here for this secondary articulation (see text).

[coronal] articulations, one of which is always a dental. However, when only one articulation must achieve complete closure, a quite distinct type

of language with three additional places is found, too, where the specialized [dorsal] articulation is a uvular not a palatal, and in all but one case, the two other additional places are labialized velars and uvulars. Both of these types are areally and perhaps genetically restricted -- the first to Australian languages and the second to North and South American or Afro-Asiatic languages. Thus they may be fairly idiosyncratic developments.

To some extent, these distributions of extra places accord with the predictions of feature geometry, though the decided preference for a specialized [dorsal] rather than [coronal] articulation among languages with just one additional place is unexpected. The frequency of the [dorsal] and [coronal] specializations are brought back to near equality, however, among languages with two additional places of articulation.

3.2 Reasons to prefer a [dorsal] rather than a [coronal] specialization

A preliminary motivation for so strongly preferring a specialized [dorsal] rather than [coronal] articulation is that a specialized [dorsal] articulation fills an empty cell in the array of place contrasts as represented by Jakobson, Fant and Halle's (195) acoustic features, sketched in table (3).

Note that the velar is specified differently in this table depending on whether the fourth place of articulation is palatal or uvular. This difference reflects an interpretation of feature values as expressing relative differences rather than absolute qualities, i.e. a velar is [+grave] in relation to a [-grave] palatal, but [-grave] in relation to a [+grave] uvular. Or in general, the fourth place in being a more extreme [dorsal] articulation in either the front or back direction along the palate determines the representation of the less extreme velar.

Stevens and Blumstein's (1978; Blumstein & Stevens, 1979) have proposed that the major places of articulation, i.e. those produced by single articulators, are distinguished by quantal differences in the distribution of energy in short-time spectra of the burst and formant transition onsets (interval including the burst and the onset of formant transitions will be referred to henceforth as the "stop release"). These spectra have a broad ([+diffuse]) energy distribution across frequencies for [labial] and [coronal] articulations, but a relatively narrow ([-diffuse] or [+compact]) energy concentration in the middle frequencies (1-3 kHz) for [dorsal] articulations. Energy distributions in these spectra for [labial] articulations tend to either

Table 3 Filling an empty slot in a system of place contrasts employing Jakobson et al.'s (1953) place features [diffuse] and [grave].

a. The major three places ([- diffuse, - grave] missing):

	p	t	k
diffuse	+	+	-
grave	+	-	+

Redundancy rule: [- diffuse] --> [+ grave]

b. Adding a palatal:

	p	t	c	k
diffuse	+	+	-	-
grave	+	-	-	+

c. Adding a uvular:

	p	t	k	q
diffuse	+	+	-	-
grave	+	-	-	+

fall as frequency increases, i.e. to exhibit the low-frequency bias represented by [+grave], or to be relatively flat, while energy distributions in these spectra for [coronal] articulations tend to rise as frequency increases, i.e. to exhibit the high-frequency bias represented by [-grave]. Distinguishing among specialized [dorsal] articulations amounts to dividing the large range of [-diffuse] (or [+compact)) spectra into high [-grave] and low [+grave] intervals.

On the other hand, specializing a [coronal] articulation would require introduction of an additional feature, e.g. [flat] to distinguish [+flat] retroflexes from [-flat] alveolars (cf. Ohala (1985) on [flat]) or to distinguish [-flat] dentals from [+flat] alveolars (note that [flat] is relativized here, too, as [grave] was to distinguish velars from palatals or uvulars).

The next two sections attempt to explain the preference for velars over palatals or uvulars among the [dorsal] articulations and the preference for alveolars over dentals or retroflexes among the [coronal] articulations in terms of how the preferred places combine articulations with mutually enhancing or integrative acoustic effects.

3.3 Perceptual reasons to prefer a velar rather than a palatal or uvular [dorsal] articulation

The question also arises as to why languages employ a velar rather than a palatal or a uvular when they have only a single [dorsal] articulation. A tentative answer to this question may be found in differences in velars vs. palatals or uvulars in: (1) whether there is any frequency bias in the energy distributions of spectra of the stop release and (2) and whether consonants at these three places coarticulate in backness with following vowels.

Regarding palatals, Blumstein (1986) presents acoustic data from Hungarian showing that the short-time spectra of palatals' stop release are compact like those of velars. The palatal short-time spectra differ from the velar in having more energy at high frequencies[11] in all vowel contexts, i.e. the palatal short-time spectra can always be considered [-grave] in contrast to the [+grave] velar short-time spectra. More precisely, the major peak in the palatal short-time spectra is higher than or equal to the frequency of F_4 in following non-rounded vowels ([i,e,a] in Blumstein's data) and higher than or equal to the frequency of F_3 in following rounded vowels ([u,o]). The short-time spectra of velar stop releases, on the other hand, are higher than or equal to a following unrounded vowel's F_3 and higher than or equal to a following rounded vowel's F_2. Furthermore, following back vowels are fronted after palatals in that their F_2's are raised substantially, while it is the consonant's backness that varies with the vowel's in velars. Keating and Lahiri (1990) present data from Czech as well as Hungarian which also show that the major peak in palatal short-time spectra has a higher frequency (which is near rounded vowels' F_3 and near unrounded vowels' F_4) than in velar short-time spectra (which is near rounded vowels' F_2 and near unrounded vowels' F_3) and that back vowels are fronted after palatals but not after velars.

The higher short-time spectra of palatals follow from the fact that their articulation at the front of the palate shortens the cavity in front of the constriction substantially compared to its length in a velar articulation. The fronting of back vowels after palatals follows from their being articulated with the tongue configuration of the palatal glide [j] combined with a post-alveolar (= palato-alveolar) blade contact (see Keating and Lahiri (1990) for x-rays and palato- and linguo-grams supporting this description of palatals' articulation in Czech and Hungarian).

Presumably, uvulars would differ in the opposite direction from velars than palatals. Articulatory data (Delattre, 1971; Ghazeli, 1977) show that the constriction is fixed at the very back of the soft palate in uvulars.

Theoretical modeling of uvular acoustics (Klatt & Stevens, 1969; see also Alwan, 1986) suggests they should lower F_2 and F_4 by lengthening the cavity in front of the constriction while leaving F_3 relatively close to its neutral value. Alwan's (1986) measurements of the acoustics of uvular consonants show that the major peak in their compact short-time spectra is close in frequency to F_2 at the onset of the following vowel before /i,a,u/ in the speech of four Arabic speakers (three from Baghdad, Iraq and the one from southern Lebanon).[12] Thus uvulars resemble velars in coarticulating in backness with following vowels. But uvulars nonetheless retain a low frequency bias (i.e. remain [+grave]) compared to ([-grave]) velars in that the major peak in the consonant's short-time spectrum stays close to the vowel's F_2 before unrounded as well as rounded vowels, rather than being closer to the vowel's F_3 as it is in velars before unrounded vowels. So uvulars don't co-articulate as much as velars for vowel backness.

In summary, palatal short-time spectra have a higher frequency bias compared to velars and uvular short-time spectra have a lower frequency bias. In addition, both palatals and uvulars vary less in backness with differences in following vowel backness than velars do. These differences in frequency bias and susceptibility to coarticulation between palatals or uvulars and velars lead to the hypothesis that the preferred [dorsal] is velar because this is the place whose major spectral peak doesn't differ in frequency much from the F_2 or F_3 of following vowels. As described above, which formant a velar's major spectral peak stays close to depends on whether the vowel is rounded. Lip rounding constricts the mouth opening and attenuates the energy radiated from the lips; this attenuation affects higher frequencies more than lower ones and thus attenuates F_3 markedly compared to F_2. On the other hand, the more open or flared mouth opening of unrounded vowels will produce a relatively more intense F_3. Therefore, it appears that a velar's major spectral peak will stay close in frequency to F_3 if it's strong enough, and otherwise F_2.

In this view, the coarticulation of velars with the backness of following vowels is deliberately controlled to ensure that the major peak in the consonant's short-time spectrum remains close to either the second or third formant of the vowel. Keating and Lahiri (1990) observe (citing data from Houde, 1968) that the tongue moves continuously along the palate in velars occurring between vowels differing in backness and may not reach the position for the next vowel until just before the consonant is released. This pattern of movement suggests that the velar reaches a back or front position appropriate to the following vowel just in time to assure that the major peak in its short-time spectrum matches that vowel's F_2 or F_3, an interpretation which accords with the view that backness coarticulation in velars is controlled rather than mechanical. As a result, the preferred [dorsal] articulation, the velar, is the one in which spectral change is minimal from the release of the consonant into at least the beginning of any following vowel. In both palatals and uvulars, the amount of spectral change between the stop burst and vowel onset will, on the other hand, be large, at least into vowels of opposite backness, compared to that obtained with velars.

Recall the general hypothesis of this paper, that articulations covary so as to produce mutually enhancing or integrative acoustic effects. By coarticulating for backness more than palatals or uvulars do, velars produce just this kind of combination of acoustic effects, i.e. the close frequency match between the major spectral peak at the stop release and the formants of the following vowel are certainly going to enhance one another and may even be integrated into a single perceptual property. This frequency match may even enhance the percept of the compact energy concentration that distinguishes [dorsal] articulations from [labial] or [coronal] articulations.

Keating and Lahiri (1990) describe another difference between velars and palatals in how the energy distribution in the spectrum changes between the stop release and the vowel onset. They note that the fronted velars that occur before front vowels may have a major spectral peak as high as that in palatals before back vowels. Thus the energy distributions in the short-time spectra of the stop release don't by themselves distinguish velars from palatals. However, even in these contexts, there is a much larger increase in energy at low than high frequencies between the stop release and vowel onset after palatals than velars, after which energy increases by about the same amount at high as low frequencies. If listeners rely on such a difference in whether energy increases equally across frequencies in identifying consonants as velar or palatal in these vowel contexts that would support the view that this place contrast depends on how the energy distribution changes in frequency between the stop release and vowel onset.[13] (A broadly similar account is offered below for the preference for alveolars over dentals or retroflexes in languages with just a single [coronal] stop.)

3.4 Perceptual reasons for preferring an alveolar rather than a dental or retroflex [coronal] articulation

It was suggested above that in order to distinguish among [coronal] articulations an additional feature was needed, beyond [diffuse] and [grave]. In the case where just two coronals contrast, this feature could be (relativized) [flat]. But there is acoustic and perceptual evidence that even adding such a feature as [flat] to represent a bias in the energy distribution in static short-time spectra of the stop release is not sufficient to distinguish dentals from both bilabials and alveolars. According to Lahiri, Gewirth and Blumstein (1984) only dynamic change in the spectral energy distribution from the stop release into the vowel reliably distinguishes these three places in languages such as Malayalam (cf. Kewley-Port (1983); Kewley-Port, Pisoni & Studdert-Kennedy (1983); and Zakia (1991) for evidence that the rate of spectral change may also be important for distinguishing velar from alveolar or bilabial places).[14]

In bilabials, either energy at low frequencies increases only slightly between the stop release and vowel onset or the increase is about the same at high as low frequencies. In dentals and alveolars, on the other hand, energy increases substantially more at low than high frequencies between the stop release and vowel onset. Alveolars are distinguished from dentals by having greater energy across the spectrum in the stop release compared to vowel onset (see also Jongman, Blumstein & Lahiri, 1984).

Since a dental may be the only representative of a [coronal] articulation in a language, as in French, Lahiri et al.'s (1984) results suggest that [coronal] articulations can only be distinguished as a class from [labial] articulations by differences in the size of an energy increase at low frequencies and not merely by differences in the static short-time spectra of the stop releases. (Lahiri et al. (1984) in fact demonstrated that differences in the amount of change at low frequencies were sufficient to classify correctly French bilabial and dental stops by place.)

A dental realization is noticeably rarer than an alveolar in languages with just a single [coronal], occurring only in 47 languages in UPSID compared to 83 for alveolars.[15] An alveolar rather than dental realization in a language with just a one [coronal] stop could be preferred because differences in static short-time spectra convey place contrasts made with independent articulators better than differences in the size of the energy change between the stop burst and vowel onset.

But rather than trying to attribute the preference for alveolar over dental (or retroflex) [coronal] to alveolars being identifiable as [coronal] just from differences in static short-time spectra between them and [labial] or [dorsal] articulations, a tack more in keeping with the general hypothesis of this paper will be developed. The mere use of the [coronal] articulator yields greater increase in energy at low than high frequencies between the stop release and vowel onset. Making this [coronal] articulation alveolar rather than dental will also ensure that the short-time spectrum of its release has a high frequency bias. This will enhance the amount of increase at low frequencies between the stop release and vowel onset and thus exaggerate the contrast between [coronal] and [labial] articulations. And again, it is likely that the two acoustic effects will integrate into a single perceptual property, i.e. a large increase in energy at low frequencies. Dental stops, on the other hand, will lack this high frequency bias in the short-time spectra of their release, and because they will therefore exhibit a smaller increase in energy at low frequencies, they will be less distinct from [labial] articulations.

What of retroflex stops, the other common specialization of a [coronal] articulation? No language in UPSID apparently as a retroflex as its sole [coronal] stop, so alveolars (or dentals) are very strongly preferred over retroflexes. Why should this be so? Stevens and Blumstein (1975) found that Hindi retroflexes are characterized by a complex pattern of spectral change between the stop release and vowel. The short-time spectra of retroflex stop releases have a major peak higher than or equal to the frequency of F_3 in a following vowel, but not higher than or equal to the frequency of the vowel's F_4. In addition, F_3 rises from vowel onset to its steady-state after a retroflex, while it falls after an alveolar. A rising F_3 is also characteristic of a velar, but a velar tends to have the major peak in its short-time spectrum below the frequency of F_3 in the following vowel onset. The reason that retroflexes are less preferred than alveolars may be that retroflexes mix an acoustic property of a [dorsal] articulation (a rising F_3 transition in the vowel) with an acoustic property of a [coronal] articulation (the major peak in the short-time spectrum of the stop release above F_3 in the following vowel onset).[16] (Ladefoged and Bhaskararao's (1983) observations of correlated cross-linguistic variability in how far back tongue contact is made and in how much the tongue tip is curled back address other aspects of Stevens and Blumstein's analysis that are irrelevant to the facts considered here.)[17]

3.5 Double articulations

Turning now to doubly articulated stops, if both articulations must achieve compete closure for the articulation to be considered double, doubly-articulated stops are relatively rare and are limited to labio-dorsals (table 1). However, if only one of the articulations must achieve complete closure, then palatalized coronals and dorsals join labialized dorsals and labio-dorsals as doubly-articulated stops. Consequently there are many more instances of such double articulations under this more liberal definition (table 2).

Only palatalization and labialization will be considered here in counting examples of doubly-articulated stops where one of the constrictions is incomplete, as the other secondary articulations recognized by

UPSID, velarization and pharyngealization, are also too rare for any statistical assessment to be meaningful (or interpretable).

Labialization is protrusion rather than closure of the lips as in labio-velar stops, but undoubtedly still the same [labial] articulator.

Traditionally, palatalization is the superimposition of the [dorsal, -back] articulation of the palatal glide /j/, i.e. a specialized use of the tongue body, on the affected consonant. In the traditional view, palatalized velars are thus indistinguishable from true palatals, and palatalized alveolars and labials are doubly articulated, dorso-coronals and dorso-labials, respectively. A number of recent proposals in the feature geometric literature (e.g. Broselow & Niyondagara, 1989; Clements, 1989; Hume, 1990, 1992) treat /j/ (as well as all the front vowels and palatal consonants) as using the [coronal] articulator instead, and thus palatalization as superimposition of a [coronal, -anterior] articulation, i.e. a specialized use of the [coronal] articulator. In this analysis, palatalized alveolars are thus indistinguishable from palatals, and palatalized velars and labials are doubly articulated, corono-dorsals and corono-labials, respectively.

Choosing between these alternatives is apparently simple: if palatalized velars never contrast with palatals, then palatalization is [dorsal, -back] but if palatalized alveolars never contrast with palatals, then it is [coronal, -anterior] instead. Unfortunately, no choice on these grounds is possible since none of the 20 palatalized velars or palatalized alveolars in UPSID occur in languages with palatal stops (17 occur in languages with just the three major places of articulation).

This is another fact not correctly predicted by feature geometry, since on either analysis of palatalization, that theory predicts that palatals should co-occur with one other of these two sets of sounds if not the other. Instead both palatalized alveolars and palatalized velars are mutually exclusive with true palatals (though not with each other, since they do co-occur). More generally, feature geometry predicts that sounds which use different major articulators, here [coronal] and [dorsal], should be able to co-occur, since it should be possible to specify sounds in a consonant inventory independently for all the major articulators.[18]

On the other hand, palatalized velars occur only 6 times in UPSID to the palatalized alveolars 12 (one language with both these sounds, Nambakaengo, is omitted from these counts as it has more than 6 places of articulation, the upper limit considered here, when stops with secondary articulations are treated as doubly-articulated). The greater rarity of the palatalized velars suggests that they might be treated as doubly articulated corono-dorsals, assuming that the possession of two articulations should make a stop rarer than one with just one place, even if that place requires that an articulator be used in a specialized way. But this assumption is suspect, since the other secondary articulation, labialization, occurs almost exclusively with stops whose primary place is [dorsal] -- 42 (.737) of 57 cases are labialized velars and 12 (.210) labialized uvulars (total .947) vs. only 3 (.053) labialized labials (2) or alveolars (1) -- and labio-dorsals must be doubly articulated rather than a specialized dorsal articulation. Since doubly articulated stops turn out not to be so rare after all, the relative rarity of palatalized velars compared to palatalized alveolars cannot be used as an argument that the former but not the latter are doubly articulated.

In light of the priority of the articulator node in the representation of any place of articulation in feature geometry, it might be expected that combinations of articulators would actually be less marked than specialized use of any one articulator. This expectation is partially borne out if stops in which only articulator achieves complete closure are included (table 2) since then there are 81 instances of such doubly-articulated stops compared to just 66 instances of a [coronal] specialization. And a specialized [dorsal] articulation is only slightly more frequent as an extra place of articulation, occurring 94 times, so a double articulation may be used about as often as a specialization of just a single articulator.

Two facts need therefore to be explained: (1) why doubly articulated stops are used so frequently in inventories with more than the major three places and (2) why labio-dorsals are the most common double articulation (some explanation for the preference for palatalized alveolars over palatalized velars would be desirable, too). Both questions turn out to have the same answer.

A labio-dorsal is the most common double articulation because these two articulations both lower the frequencies of F_2 and/or F_3, in particular they will lower F_2 more than either a [labial] or [dorsal] articulation alone (see Fant, 1960; Ohala & Lorentz, 1977, for relevant discussion). In terms of Jakobson et al.'s (1952) features, these two articulations combine [+flat] with [+grave].

No other combination of articulations (with the possible exception of a palatalized alveolar, see immediately below) produces acoustic effects which enhance one another in the way the effects of [labial] and [dorsal] articulations do. In fact, at least some other combinations yield combinations of acoustic effects that confuse the listener about the consonant's place. For example, combining a palatal and a labial articulation raises F_2

in an adjacent vowel sufficiently that palatalized labials frequently become alveolars or palato-alveolars (Ohala, 1978). Palatalized velars become palatals or palato-alveolars for similar reasons.

So the high frequency of labio-dorsals among doubly articulated stops and relatively high frequency of double articulations compared to a specialized use of a single articulator are both a reflection of the mutual enhancement of F_2 and F_3 lowering accomplished by combining [labial] and [dorsal] constrictions.

The relatively high frequency of the combination of a palatal articulation with a [coronal] one in palatalized alveolars may have a similar explanation, i.e. palatalizing an alveolar with add [+sharp] to the [-grave] specification of the alveolar, increasing the bias toward a (very) high frequency energy concentration. The low frequency of these dorso-coronals, however, compared to the much more frequent labio-dorsals may be a reflection of their similarity to another species, palato-alveolars, which are discussed extensively in the next section.

3.6 Summary

In this section, the frequency of occurrence has been assessed of stops made with specialized [coronal] and [dorsal] articulations, as well as of doubly articulated stops. Both systemic and perceptual explanations were developed for the clear preference for a specialized [dorsal] over a specialized [coronal] articulation, for the preferences for alveolar and velar stops in languages which have just a single [coronal] or [dorsal] stop, respectively, and for labio-dorsal doubly articulated stops. The perceptual explanations took the same form in all cases: the preferred or more frequent place is one produced by combining articulations whose acoustic effects mutually enhance one another. Indeed, the perceptual interaction among the acoustic effects of these combined articulations may go beyond enhancement to actual integration into a single perceptual property, as many of the combined articulations affect the acoustics in very similar ways.

4 Affricates as an alternative means of increasing place contrasts for stops

4.1 Affricates are a variety of stop

In the discussion of double articulations above, a different picture of their relative frequency emerged when what counts as a double articulation was defined liberally to include any stop with two constrictions even when one did not achieve complete closure. Another liberalization, including affricates among the stops, changes the perspective developed so far even more dramatically. Including affricates among the stops also suggests that the principles that determine what contrasts are common in languages with larger numbers of place contrasts are quite different from any derived from feature geometry.

In fact, once affricates are considered as a species of stop, a specialized [coronal] articulation grossly outnumbers any specialized [dorsal] articulation, because the only places of articulation where affricates are at all common are palato-alveolar or alveolar (0.56 of the affricates in UPSID are palato-alveolar and 0.32 are alveolar, together 0.88 of all the affricates). But before we can consider how [coronal] affricates contribute to the expanding the number of place contrasts in a language, it's necessary to provide positive arguments for considering them as stops.

Affricates resemble stops, of course, in that they are produced with a complete interruption of air flow out of the mouth. But sibilant affricates resemble stops in another way, in that they have a noticeably voiced:voiceless ratio than the corresponding sibilant fricatives (see above for related discussion). Voicing is actually disfavored in sibilant fricatives, where the voiced:voiceless ratio is only 0.34, compared to 0.62 for non--sibilant fricatives. Affricates reverse this asymmetry, showing voiced:voiceless ratios of 0.47 for sibilant affricates vs. 0.39 for non-sibilant ones, which makes sibilant affricates resemble stops, whose voiced:voiceless ratio is 0.62. The following discussion explores this resemblance of sibilant affricates to stops further and introduces a more sophisticated means of assessing statistically the co-occurrence dependencies among categorical properties of speech sounds, log-linear models (Fienberg, 1980, especially chapters 3 and 4). As log-linear models of such dependencies are used extensively in the remainder of this paper, their use will be outlined below (see the appendix for details).

Loglinear models generalize the familiar X^2 analysis of observed vs. expected frequencies of data classified by two categorical variables to data classified by more than two such variables. In place of calculating the X^2 statistic, a similar statistic, G^2, is calculated instead, which also gets large when observed frequencies differ from expected frequencies. In moving to more than just two classificatory variables, a series of models are built reflecting a hierarchy of assumptions: first about whether the variables affect the data independently of one another and then about possible interactions among the variables in how they affect the data. Each model of the data yields different expected frequencies, and thus different G^2 statistics. The smaller G^2 is for a particular model, the less the discrepancy between the expected frequencies calculated in that model and the observed frequencies, and the better the "fit" to the data. Model calculation continues until G^2 is small and no longer reduced significantly by adding more interactions to the current model. That model is then used to interpret the independence/dependence of the variables' effects on the data.

The first set of log-linear models that will be considered here assess the dependencies between the variables: voicing, sibilance, and affrication,[19] in affricates and fricatives. Table 4 tabulates the frequency of affricates and fricatives, cross-classified by voicing and sibilance.

Table 4 O(bserved) vs. E(xpected) frequencies for models [1,2,3] (all variables independent), and [13,12], in which sibilance interacts with both the affricate:fricative and voiced:voiceless contrasts.

(3)		Affricate		Fricative		Model	G^2
(2)		Vcd	Vls	Vcd	Vls		
Sibilant	O:	119	252	157	459		
	E:	81.1	178.1	227.8	500.0	[1,2,3]	249.7
	E:	103.8	267.3	172.3	443.8	[13,12]	6.0
(1)							
Non-sibilant	O:	7	18	189	307		
	E:	42.8	94.0	120.3	263.9	[1,2,3]	
	E:	9.4	15.6	186.6	309.4	[13,12]	

G^2 for the model with complete independence of all the variables, [1,2,3], is 249.7, a very poor fit. G^2 is reduced substantially to 20.6 (which is still a significant value), by adding the interaction between sibilance and affrication [13]. By adding to that model the interaction between sibilance and voicing [12] it is reduced further to 6.0, just over the significance criterion of 5.99, $p < 0.05$, for a model with two degrees of freedom. Since one variable is independent of at least one of the others in each of these interactions, i.e. variable 2, voicing, is independent of variable 3, affrication, it's possible to collapse across voicing or affrication to assess these interactions in 2x2 contingency tables (Fienberg 1980: 48-51). Collapsing table 4 in this way, into tables 5 and 6, simplifies the interpretation of the dependencies among these variables. Collapsing across voicing (table 5) shows that there are many more sibilant affricates and many fewer non-sibilant affricates than expected, with complementary differences for the fricatives.

Table 5 Observed and Expected frequencies for sibilant and non-sibilant affricates, collapsing across voicing.

(3)		Affricate	Fricative	Total
Sibilant	O:	371	616	987
	E:	259.2	727.8	
(1)				
Non-sibilant	O:	25	496	521
	E:	136.8	384.2	
Total		396	1112	1508

$G^2(1) = 229.1$, $p < 0.001$

Collapsing across affrication (table 6) shows that there are fewer voiced sibilants and more voiced non-sibilants than expected; for voiceless sounds, the differences between expected and observed values are rather small. (The discrepancies between observed and expected frequencies are smaller in table 6 than table 5, but nonetheless still highly significant.)

When treated separately like this, the two interactions show that sibilant affricates are more common than expected and voiced sibilants are less common that expected. Given these tendencies, the larger voiced:voiceless proportion in sibilant than non-sibilant affricates or than in sibilant fricatives becomes that much more striking. Sibilant affricates are clearly more like stops than sibilant fricatives in how voicing co-occurs with

Table 6 Observed vs. Expected frequencies for voiced and voiceless sibilants and non-sibilants collapsing across affrication.

(2)		Voiced	Voiceless	Total
Sibilant	O:	276	711	987
	E:	308.9	723.9	
(1)				
Non-sibilant	O:	196	395	521
	E:	163.1	382.1	
Total		472	1106	1508

$G^2(1) = 10.5,\ p < 0.001$

them. Now that the stop-like character of sibilant affricates is established we can return to their contribution to increasing the number of places of articulation for stops in a language.

4.2 The patterning of sibilant palato-alveolar affricates

4.2.1 Introduction

The following discussion focuses on the patterning of the most frequent affricate, the sibilant palato-alveolar. It will first be shown that palato-alveolar affricates tend not to occur in languages which have a non-sibilant palatal stop or affricate or a palato-alveolar stop. This suggests that these sounds could all be considered equivalent means of increasing place contrasts in a stop inventory. It will next to be shown that the occurrence of a palato-alveolar affricate in a language depends on the occurrence of both a corresponding palato-alveolar fricative and another sibilant affricate. These three sounds, together with the ubiquitous dental/alveolar fricative /s/, tend to appear as a complex of affrication and place contrasts, what will be called the "sibilant" contrast block.

4.2.2 Palato-alveolar affricates vs. non-sibilant palatal non-continuants

The first question is whether palato-alveolar affricates occur freely as an additional place in a stop inventory. If so, then this [-anterior] specialization of the [coronal] articulation outnumbers any other means of increasing place contrasts beyond the major three, since palato-alveolar affricates occur in over 170 of the 317 languages in UPSID.

However, sibilant palato-alveolar affricates are not completely free in their occurrence in that they tend not to co-occur with palatal stops. This failure to co-occur raises the possibility that sibilant palato-alveolar affricates and palatal stops are phonetically distinct realizations of a non-continuant with the same phonological specification for place (at least in the languages where they don't co-occur). The markedly greater frequency of the sibilant palato-alveolar affricate realization of this non-continuant suggests that the affricate has phonetic properties that mak it a better realization of this place than palatal stops. Treating these sounds as phonetically but not phonologically distinct also suggests that the phonological representation of palatal stops might not be the [-back, dorsal] specification assumed above, given that the sibilant palato-alveolar affricates are generally thought to be a [-anterior] specialization of a [coronal] articulation. The proper phonological representation of the putative single place of these sounds will be taken up after the co-occurrence dependencies between sibilant palato-alveolar affricates and palatal stops are illustrated.

In examining the co-occurrence of palato-alveolar affricates with palatal stops, the 10 languages with palatal affricates and the 8 languages with palato-alveolar stops have been lumped together with the 55 languages with palatal stops; these three kinds of [-anterior, coronal] non-continuants. This amounts to distinguishing phonetically between the non-sibilant, i.e. [-strident], realization of the palatal stops and affricates and the palato-alveolar stops and the sibilant, i.e. [+strident], realization of the palato-alveolar affricates (henceforth the [-strident] sounds will be referred to as "non-sibilant palatals" and the [+strident] sounds as "sibilant palato-alveolars"). Support for collapsing the [-strident] sounds into a single phonetic class comes from the fact that the non-sibilant palatals are nearly entirely mutually exclusive. There is just one language in UPSID, Komi, with both palatal stops and affricates, but there are no languages with both palatal and palato-alveolar stops nor any with both palatal affricates and palato--alveolar stops. (Leaving the palato-alveolar stops out of the comparisons below does not change the results.) Table 7 cross-classifies the languages in UPSID in terms of whether they have such a broadly-defined non--sibilant palatal (variable 3) or a sibilant palato-alveolar (variable 2), and by the number of other places of articulation distinguished (variable 1).

Languages are categorized by variable 1 in table 7 into those languages which have no other additional places of articulation beyond those contributed by a sibilant palato-alveolar or a non-sibilant palatal vs. those which have other additional places besides either of these. This variable thus allows a comparison of whether and how sibilant palato-alveolars or non-sibilant palatals augment a stop inventory in languages with and without yet other specialized uses of articulators, or the combination of two articulators.

Table 7 Observed vs. Expected numbers of lgs. for sibilant palato-alveolars and non-sibilant palatals in languages with 0 vs. more than 0 additional places of articulation for stops.

(3) Palatals		Yes	Yes	No	No	Model	G^2
(2) Palato-alveolar affricates		Yes	No	Yes	No		
0	O:	21	30	98	70		
	E:	27.1	23.4	90.4	78.2	[1,2,3]	22.9
	E:	16.6	33.9	100.9	67.7	[23]	6.4
(1) Additional places							
> 0	O:	3	19	48	28		
	E:	12.1	10.5	40.5	35.0	[1,2,3]	
	E:	7.4	15.2	45.1	30.3	[23]	

G^2 for the model [1,2,3], where all the variables are independent, is 22.9, a relatively poor fit, but adding the interaction between non-sibilant palatals and sibilant palato-alveolars [23], reduces G^2 significantly to just 6.4, at 3 d.f. $0.10 > p > .05$ (neither of the other single interaction models reduce G^2 significantly, nor do adding more interactions to [23]). This interaction is significant because there are fewer than the expected number of languages with both sibilant palato-alveolars and non-sibilant palatals (first column), and more than the expected number with just one of the other of these two categories (second and third columns).

That neither interaction involving the presence or absence of additional places of articulation improves the fit shows that the likelihood of either non-sibilant palatals or sibilant palato-alveolars occurring in a language does not depend on whether the language has additional places of articulation beyond the major three. This also means that they tend not to co-occur in languages which otherwise just have the major three places of articulation as well as in languages with other additional places. Collapsing across the variable of additional places (table 8) shows that in languages with non-sibilant palatals, sibilant palato-alveolars are less common than expected, while in those without they are more common. This result supports the hypothesis that non-sibilant palatals and sibilant palato-alveolars are phonetically but not phonological distinct sounds.

Table 8 Observed vs. Expected numbers of languages with non-sibilant palatals and sibilant palato-alveolars, collapsing across the presence of additional places of articulation.

			(3) Non-sibilant palatals		
			Yes	No	Total
(2) Sibilant palato-alveolars	Yes	O:	24	146	170
		E:	39.2	130.9	
	No	O:	39	98	147
		E:	33.9	113.2	
	Total		73	244	317

$G^2(1) = 8.6$, $p < 0.001$

Two questions are raised by the distributions laid out here: why do non-sibilant palatals and sibilant palato-alveolars tend not to co-occur and why do sibilant palato-alveolars occur three times as often as non-sibilant palatals.

It was suggested above that the answer to the first question is that these two kinds of sounds are merely phonetically distinct realizations of the same single phonological place.[20] What then is the phonological representation of this place? The greater frequency of the palato-alveolar than palatal realization would support the choice of the [coronal] rather than [dorsal] articulator (cf. Clements 1976, 1989; Broselow & Niyondagara, 1989; Hume 1990, 1992). This place would furthermore be [-anterior] to ensure its post-alveolar realization.

The increase in the intensity of the frication noise that results from making these post-alveolar stops sibilant ([+strident]) enhances the mid-range energy concentration in their release spectra, i.e. it accentuates the contrast between these [-diffuse] post-alveolars and the [+diffuse] alveolars. Sibilance may, furthermore, be more characteristic of palato-alveolars than palatals because their more anterior articulation ensures that the air jet strikes the teeth immediately after passing through the constriction (Shadle 1985). However, the fact that non-sibilant palato-alveolar stops occur (even if rarely) shows that the location of the constriction does not produce sibilance by itself. Sibilance also cannot be a product of the [coronal] articulation, since the interdental fricatives [θ,ð] are [coronal] without being sibilant. Sibilance instead requires specific adjustments of the shape of the channel through which the air rushes to create the desired high intensity noise. The preference for sibilant palato-alveolars apparently follows from deliberate adjustments of the location and shape of the constriction designed to enhance the [diffuse] contrast. The next section shows that once sibilance is used to enhance the contrast between alveolar and post-alveolar coronals, a substantial variety of place and manner contrasts is likely to appear among sibilants.

4.2.3 Conditions on the occurrence of sibilant palato-alveolar affricates

The occurrence of a sibilant palato-alveolar affricate turns out not to be entirely unconstrained. Instead, a sibilant palato-alveolar affricate is more likely to occur in a language if that language also has a sibilant palato-alveolar fricative and another sibilant affricate (in the vast majority of cases, a dental/alveolar affricate, i.e. /ts/, though occasionally a retroflex affricate, i.e. /ʈʂ/). (The relationship between inventory size and the occurrence of a sibilant palato-alveolar affricate is taken up in the next section.)

Table 9 below cross-classifies the languages in UPSID by the presence of a sibilant palato-alveolar affricate, a sibilant palato-alveolar fricative, and another sibilant affricate.

The very large G^2 of 105.4 for model [1,2,3], in which all the variables are independent, indicates a very poor fit. Model [23], the interaction between a palato-alveolar affricate and fricative, reduces G^2 the most,

Table 9 Observed vs. Expected numbers of languages with a [+strident] palato-alveolar affricate /tʃ/, the corresponding palato-alveolar fricative /ʃ/, and another sibilant affricate, most often /ts/.

(3) Palato-alveolar fricatives		Yes	Yes	No	No	Model	G^2
(2) Palato-alveolar affricates		Yes	No	Yes	No		
(1) Another sibilant affricate							
Yes	O:	63	9	8	33		
	E:	29.0	24.1	32.7	27.2	[1,2,3]	105.4
	E:	42.4	10.7	19.3	40.6	[23]	29.3
	E:	57.5	14.5	13.2	27.8	[13,23]	9.4
No	O:	56	21	46	81		
	E:	52.3	43.6	59.0	49.1	[1,2,3]	
	E:	76.6	19.2	34.8	73.4	[23]	
	E:	61.5	15.5	40.8	86.2	[13,23]	

to 29.3, of all the single two-way interaction models. The fit of this model can in turn be improved the most by adding the interaction between another sibilant affricate and a palato-alveolar fricative [13], which yields a G^2 of 9.4. This is still a significant value ($p < 0.01$, d.f. = 2), but no model with additional interactions has a significantly lower G^2. Thus, the model in which the presence of /ts/ [13] or /ʃ/ [23] interacts with the presence of /tʃ/ fits the data best.

Interpreting this result can be simplified substantially as before by reducing this 2x2x2 table to a pair of 2x2 tables because one of the variables, [2], a palato-alveolar affricate, remains independent of another, [1], another sibilant affricate, in the 2x2x2 model that fits the data best. Collapsing across these variables in turn, yields in the 2x2 tables in (10); (10a) collapses across variable [1], another sibilant affricate, and (10b) across [2], a palato-alveolar affricate.

Table 10 Observed vs. Expected numbers of languages for (a) Another sibilant affricate (1) vs. Palato-alveolar fricatives (3) and (b) Palato-alveolar affricates (2) vs. Palato-alveolar fricatives (3).

(a)

(3) /ʃ/		Yes	No	Total
(1) /ts/				
Yes	O:	72	41	113
	E:	53.1	56.3	
No	O:	77	127	194
	E:	91.2	22.3	
Total		149	168	317

$G^2 = 45.4$, $p < 0.001$

(b)

(3) /ʃ/	Yes	No	Total
(2) /tʃ/			
Yes	119	54	173
	81.3	91.7	
No	30	114	144
	67.7	76.3	
	149	168	317

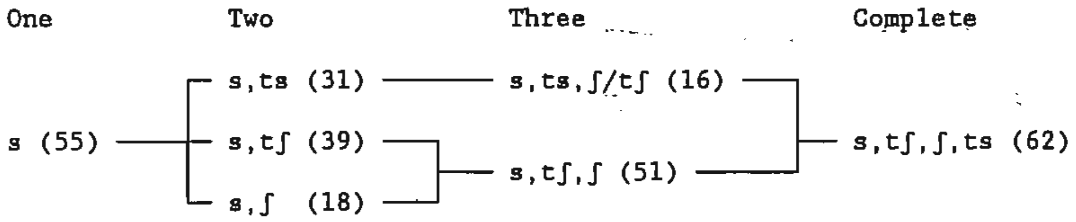
$G^2 = 76.1$, $p < 0.001$

The highly significant G^2's in both these 2x2 tables show that palato-alveolar affricates and other sibilant affricates each co-occur more often with a palato-alveolar fricative than expected, and that all three of these sibilants are less likely than expected to occur alone. More simply, these three types of sibilants are more likely to occur together or to all be absent than expected from their individual frequencies.

Furthermore, /s/ is extremely common in the languages in UPSID (occurring in 0.86 of the sample). /s/ is also rarely absent in languages that have any other sibilant, and in the 81 languages which have neither a palato-alveolar fricative or affricate nor another sibilant affricate, /s/ is also absent in 26, nearly 3 times as often as in any of the languages with another sibilant.

These co-occurrence dependencies show that languages are statistically likely to be of one of three types with respect to what sibilant inventories they have; either they will have: (1) no sibilants, (2) a single sibilant, which is nearly always /s/, or (3) a complete set of sibilants consisting of [+] and [-anterior] sibilant fricatives and affricates. Table 11 lists the observed frequencies in UPSID of the possible sets of sibilants for languages with at least one sibilant:

Table 11 Frequencies of various numbers of sibilants.



This table shows that a language is likely to have either just a single sibilant, in which case it's [+anterior, +continuant, coronal], i.e. /s/, or to have four sibilants, contrasting for place (in the feature [anterior]) and continuancy. i.e. /s, ʃ, ts, tʃ/. The filled sibilant inventory is an example of what was called a "contrast block" above, a set of segments that use all possible combinations of values for a set of distinctive features, here [anterior] and [continuant].

Recall that the hypothesis represented by the notion of contrast blocks is that languages with larger inventories will differ from those with smaller ones in what contrast blocks they employ, rather than simply in the identity and number of individual segments. There are 55 languages in UPSID with the minimal sibilant block and 62 with the filled contrast block.

But what about the languages which only partially fill this contrast block? 67 languages have just two of the sibilants in addition to /s/, and 88 have one other besides /s/. Though a substantial number of languages clearly have only partial contrast blocks, these partial blocks are less frequent than expected from the individual frequencies of their members.

The asymmetries in the frequencies of the members of these partial contrast blocks suggest a likely evolutionary course for filling the sibilant contrast block a segment at a time. The occurrence of /tʃ/ does not apparently depend on the prior occurrence of the corresponding fricative /ʃ/, and either of the affricates are more likely to occur as the second sibilant than /ʃ/, and the presence of /ʃ/ implies the presence of /tʃ/.

One can still ask what further arguments there are for the sibilant contrast block as more than the simple sum of its constituent segments. The first argument is simply that there are more than the expected number of languages with all three sibilants in addition to /s/, and more than expected number of languages with just /s/. The second is that there are also markedly fewer than the expected number of languages with just any one or two of the sibilants in addition to /s/ (see table 9 above); this is particularly true for languages with another sibilant affricate and one of the palato-alveolars, but it's also true for languages with just one of the affricates, either a palato-alveolar or another sibilant.

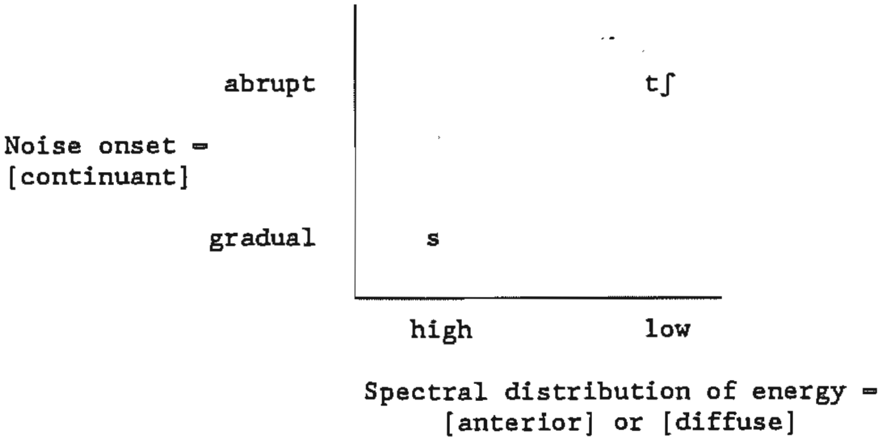
One can also ask what sibilant is likely to be added to an inventory which just has /s/. The answer is that it's almost certainly /tʃ/, since this is the most common sibilant after /s/ and the one that occurs most frequently in the partial contrast blocks. But why should it be /tʃ/ rather than one of the other sibilants, i.e. why should the complexity of the sibilant inventory increase along two dimensions rather than just one (in defiance of expectations from markedness theory)?

The answer may lie in Lindblom's (1986, 1990, see also Lindblom & Maddieson, 1988, and Ohala, 1979) theory of adaptive dispersion, the idea that languages combine minimization of articulatory cost (see also Lindblom, 1983) with maximization of the perceptual distance between contrasting sounds. Lindblom's proposals have been tested for both vowels (Lindblom, 1986) and consonants (Lindblom & Maddieson, 1988). That all the sounds involved are sibilants constrains the dimensions that can be manipulated for the purposes of conveying contrasts, i.e. the sounds must remain [coronal] and [+strident] -- this is the minimization of articulatory cost.[21] The remaining dimensions are precisely those that distinguish affricates from fricatives, i.e. the abruptness of noise onset, and [-anterior] from [+anterior] coronals, i.e. the spectral distribution of energy in the noise. /tʃ/ differs from /s/ along both these dimensions.

And since it does, /tʃ/ opens up the cells for the sounds which differ from it and /s/ in just [continuant] or [anterior] (or [diffuse]), /ʃ/ and /ts/, when it enters the inventory, i.e. in being a [-anterior, -continuant], it opens up the cells for a [-anterior, +continuant] and a [+anterior, -continuant].

The mutual dependencies just demonstrated between the presence of a palato-alveolar affricate and other sibilants, both fricatives and affricates, show that adding a palato-alveolar affricate to an inventory does not just add palato-alveolar to the place contrasts in the stop series, but also tends to add further place contrasts, between alveolar and palato-alveolar affricates, and manner contrasts, between the palato-alveolar affricates and the corresponding fricatives. That is, a palato-alveolar affricate tends to occur as part of a contrast block consisting of other sibilant obstruents.

Table 12 Dimensions of acoustic contrast between sibilants: abruptness of noise onset (= [continuant]) and spectral distribution of energy in the noise (= [anterior] or [diffuse]).



4.2.4 The relationship between the sibilant contrast block and inventory size

But before accepting the notion that /ʃ, tʃ, ts/ tend to appear along with /s/ as a block in a language, the possibility must be considered that these sounds tend to occur individually in inventories above a certain size, and that so long as an inventory is large enough, it's likely that more than one of these sounds will occur in it (cf. Lindblom & Maddieson, 1988). After all, the individual frequencies of occurrence of these sounds are all fairly high and not grossly different, which predicts that their co-occurrence should not be infrequent. Although attributing the frequency of co-occurrence of these sounds to sufficient inventory size is incompatible with the fact that a model with two two-way interactions fits the data best (cf. table 9), it would still be useful to examine the effects of inventory size on their occurrence.

The effects of inventory size can be gauged by comparing the frequency of co-occurrence of the sounds inside the sibilant contract block with another sound from outside the block whose overall frequency is similar. Labial fricatives (lumping bilabial and labio-dental places) were chosen for this purpose, since their frequency of occurrence 180/317 = 0.57 is of roughly the same magnitude as that of palato-alveolar affricates (170/317 = 0.54) and fricatives (149/317 =0.47).

If the occurrence of either palato-alveolar affricates or fricatives is a large inventory effect, then their occurrence shouldn't depend on the presence or absence of labial fricatives. On the other hand, if increasing inventory size also systematically activates unused combinations of distinctive feature values of the sort captured in the notion of a contrast block (see Lindblom & Maddieson, 1988, for an argument of this sort), then more than the expected number of languages would have both labial and palato-alveolar fricatives, since labial and palato-alveolar fricatives could comprise a contrast block which appears when a language expands its fricative inventory beyond `/s/`. But no more than the expected number would have both labial fricatives and palato-alveolar affricates.

Table 13 below shows that `/ʃ/` is significantly more likely to occur in a language which also has `/f/` (or `/ɸ/`) and significantly more than the expected number of languages lack both fricatives. Table 14 shows, on the other hand, that the occurrence of `/tʃ/` does not depend on the occurrence of `/f/` (or `/ɸ/`). The first result suggests that once a place contrast arises among the fricatives, i.e. in a language with more than just `/s/`, a fricative contrast block `/{f,ɸ},s,ʃ/` with two fricatives contrasting for place with `/s/` is likely to appear (see below for a modification of this result). The second result shows that `/tʃ/` doesn't just occur in a language whose inventory size exceeds some limit (cf. Lindblom & Maddieson, 1988) and thus does not tend to co-occur with the other sounds in the sibilant contrast block just because these sounds are all common in inventories above a certain size. (Two comparisons are made in each table, one in which bilabial fricatives are included with labio-dentals in the counts of labial fricatives and one where they're not; the results don't differ. Though these are 2x2 tables, G^2 tests are still used for consistency.)

4.2.5 The fricative contrast block and its relationship to the sibilant contrast block

The fricative contrast block identified in the previous section turns out to be incomplete, in that it should include a [dorsal] fricative as well. Table 15 illustrates co-occurrence frequencies among [coronal] (other than `/s/`), [labial], and [dorsal] fricatives.[22]

G^2 is quite large, 54.8, for the model where all the variables are independent, but can reduced significantly twice: first to 16.7 by adding the interaction between another [coronal] fricative and a [dorsal] fricative [13], and then to 1.5 (n.s. for 2 d.f.) by adding the interaction between another [coronal] fricative and a [labial] fricative [12] (recall tables 13 and 14).

Table 13 Observed vs. expected numbers of languages for palato-alveolar fricatives and labial fricatives, when the latter are either narrowly (left) or broadly (right) construed.

			Labio-dental alone			Bilabial and Labio-dental		
Labial fricatives			Yes	No	Total	Yes	No	Total
Palato-alveolar fricatives	Yes	O:	82	63	145	100	45	145
		E:	67.7	77.3		82.8	62.2	
	No	O:	66	106	172	81	91	172
		E:	80.3	91.7		98.2	73.8	
	Total		148	169	317	181	136	317
			$G^2(1) = 10.5$, $p < 0.001$			$G^2(1) = 15.5$, $p < 0.001$		

Table 14 Observed vs. Expected numbers of languages for palato-alveolar affricates and labial fricatives (narrowly and broadly construed).

			Labio-dental alone			Bilabial and Labio-dental		
Labial fricatives			Yes	No	Total	Yes	No	Total
Palato-alveolar affricates	Yes	O:	82	89	171	105	66	171
		E:	79.3	91.7		97.1	73.9	
	No	O:	65	81	146	75	71	146
		E:	67.7	78.3		82.9	63.1	
	Total		147	170	317	180	137	317
			$G^2(1) = 0.4$, $p > 0.10$			$G^2(1) = 3.2$, $0.10 > p > 0.05$		

Since again at least one variable is independent of one of those involved in

Table 15 Observed vs. Expected numbers of languages with (1) [coronal] fricatives (other than /s/) /ʃ,θ,ʂ,ɬ/, (2) [labial] fricatives /f,ɸ/, and (3) [dorsal] fricatives /ç,x,χ/.

(3) Dorsal fricative		Yes	Yes	No	No	Model	G^2
(2) Labial fricative		Yes	No	Yes	No		
(1) Coronal fricative							
Yes	O:	64	30	58	30		
	E:	39.4	28.9	65.6	48.1	[1,2,3]	54.8
	E:	54.3	39.7	50.8	37.2	[13]	16.7
	E:	63.0	31.0	59.0	29.0	[12,13]	1.5
No	O:	14	11	47	63		
	E:	29.3	21.4	48.7	35.6	[1,2,3]	
	E:	14.4	10.6	63.5	46.5	[13]	
	E:	11.3	13.7	49.7	60.3	[12,13]	

these interactions, i.e. the presence of a [labial] fricative [2] is independent of the presence of a [dorsal] fricative [3], this three-way table can be collapsed to a pair of two-way tables (16a,b) to aid its interpretation.

Table 16a, in which variable [2], a labial fricative, is collapsed, shows that there a significantly more languages than expected with both [coronal] fricative (other than /s/) and a [dorsal] fricative and significantly fewer than expected with just one of these fricatives. Table 16b shows an even stronger statistical tendency for a [labial] and a [dorsal] fricative to occur together or to both be absent than for either to occur alone.

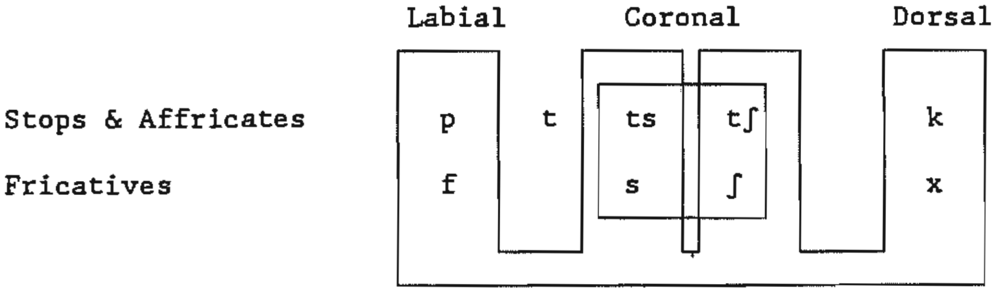
This expanded fricative place contrast block, consisting of /f,s,ʃ,x/ (/s/ is included because of its exceptionally high frequency) interlocks with the sibilant contrast block in that by far the most common [coronal] fricatives, /s/ and /ʃ/, belong to both blocks. The interlock is more than a matter of sharing members, since it's already been shown that the occurrence of /ʃ/ depends on the occurrence of /tʃ/. As argued above, the presence of /tʃ/ increases the likelihood that /ʃ/ will also occur (it also

Table 16 Observed and Expected frequencies of co-occurrence for (a) [labial] vs. [dorsal] and (b) [coronal] vs. [dorsal] fricatives.

(3) /x/		Yes	No		Yes	No		(3) /x/
(1) /ʃ/								(2) /f/
				Total			Total	
Yes	O:	78	105	183	94	88	182	Yes
	E:	68.7	114.3		68.3	113.7		
No	O:	41	93	134	25	110	135	No
	E:	50.3	83.7		50.7	84.3		
Total		119	198	317	119	198	317	
		$G^2 = 4.8,\ p < 0.05$			$G^2 = 38.1,\ p < 0.001$			

makes /ts/ more likely to co-occur). Once /tʃ/ and /ʃ/ contrast in a language, slots apparently open up for fricatives contrasting with stops made with the [labial] and [dorsal] articulators. The interlock of the full fricative and sibilant contrast blocks is laid out in table 17.

Table 17 The interlock of the sibilant and fricative contrast blocks.



What the diagram in table 17 implies is that the organization of obstruents by place and manner does not arise because these are primitive organizing properties for obstruents but instead out of fitting together of contrast blocks. In particular, the addition of /ʃ/ to the sibilant contrast block provides the structure for the stop:fricative contrasts made with the other articulators. As classificatory properties, place and manner are thus emergent rather than organizing properties; but once present, they may provide a stabilizing feedback on the structure of the inventory.

4.2.6 Summary

This section demonstrated first that sibilant palato-alveolars and non-sibilant palatals are likely just phonetically distinct realizations of a post-alveolar stop. It also argued that the markedly higher frequency of a sibilant realization of a stop at this place of articulation stems from how sibilance enhances the [diffuse] contrast between alveolars and post-alveolars. It was then shown that sibilant palato-alveolars do not occur freely in languages, but instead tend to occur as part of a block of sibilants contrasting for place and manner, i.e. /s, ʃ, ts, tʃ/. Finally, since a member of this sibilant contrast block, /ʃ/, tends to co-occur with [labial] and [dorsal] fricatives, the sibilant contrast block interlocks with a fricative contrast block.

5 Concluding remarks

When a narrow definition of stops, which excludes affricates, is used in examining the differences between languages with small and large numbers of place contrasts for stops, then languages with a larger number of contrasts specialize in their use of the tongue tip-blade and tongue body articulators, [coronal] and [dorsal], or combine [labial] and [dorsal] articulations in doubly articulated stops.

However, once affricates are included, the story changes. The conditions on the occurrence of the very common sibilant palato-alveolar affricates suggest that inventories of obstruents, (continuants as well as non-continuants) are composed of contrast blocks. Contrast blocks are sets of sounds selected so as to minimize articulatory cost and maximize perceptual distance at the same time, a combination of constraints captured by Lindblom's theory of adaptive dispersion.

Two interlocking contrast blocks were identified in the languages of UPSID, a sibilant block consisting of alveolar and post-alveolar [coronal] fricatives and affricates and a fricative block consisting of [labial], [coronal], and [dorsal] fricatives. Larger inventories differ from smaller ones then not just in having more, increasingly marked segments but in having more contrast blocks. The significantly greater frequency of inventories with complete contrast blocks suggests finally that the markedness of an inventory is not a monotonic function of its size, but instead a function of whether its contrast blocks are complete.

6 Appendix on loglinear models

In loglinear models of categorical data, a hierarchy of models is built up, representing different assumptions about co-occurrence dependencies between sets of sounds categorized by the variables in the analysis. Consider tables 4 and 9, which represent the frequency of occurrence in the UPSID sample of sounds varying in (1) sibilance, (2) voicing, and (3) affrication (table 4) and the frequency of occurrence of languages with (1) another sibilant affricate, (2) palato-alveolar affricates, and (3) palato-alveolar fricatives (table 9). In tables like 4, where it is the number of occurrences of a sound of a particular type that is being assessed, each cell in the table represents the number of times a sound occurs with the combination of properties indicated in the row and column labels, e.g. the upper lefthand cell in table 4 indicates that a voiced sibilant affricate is observed 119 times in the UPSID sample. In tables like 9, on the other hand, each cell containing an o(bserved) frequency represents the number of languages which do (Yes) or do not (No) possess the indicated combination of sounds, e.g. the value 63 in the upper lefthand corner is the number of languages which have another sibilant affricate, a palato-alveolar affricate, and a palato-alveolar fricative, i.e. the number of languages for which all three variables take the value "Yes." (Note that the variables are either privative or more rarely binary.)

Different assumptions about the co-occurrences among each of these sets of sounds are represented in a hierarchical series of models, each of which yields a set E(xpected) frequencies. These models are assessed by comparing the expected with the observed frequencies, and large discrepancies between observed and expected frequencies -- revealed by large G^2's -- indicate that the assumptions about co-occurrence in a particular model don't fit the observed patterns of co-occurrence very well.

For the data in table 4, the hierarchy begins with a model in which all the variables, here properties of sounds, are assumed to be independent of one another, meaning that a sound having one of these properties is assumed not to influence it's having one of the others. (In those tables, e.g. table 9, where observed frequencies are numbers of languages, the model assuming complete independence assumes that the occurrence of one kind of sound in a language doesn't depend on the occurrence of the other two types of sounds.) In such a model, the frequency with which properties co-occur in a sound is equal to the product of the frequency of occurrence of each individual property. (Similarly, the number of languages having a particular set of sounds should be equal to the product of the frequencies of occurrence of each sound in the set.)

Using 1, 2, and 3 to indicate the three variables that will be employed in the various loglinear analyses carried out here, the model of complete independence can be represented as [1,2,3]. The lower models in the hierarchy include interactions between categories, i.e. dependencies between the occurrence of one class of sounds and another: first, single pairwise interactions are considered one by one in addition to the three independent variables (models [12], [13], or [23]); second, models are considered in which each of two variables interact with the third, but they are independent of one another (e.g. [13,23], where variables 1 and 2 interact with variable 3 but not with each other, etc.); third, models are considered in which all three variables interact in pairs with each other ([12,13,23]); and finally, the model is considered in which the three-way interaction, as well as all the pairwise interactions are significant ([123]).

A model with a single pair-wise interaction such as [12] assumes that occurrence of the sounds categorized by variable 1 is dependent on the occurrence of sounds categorized by variable 2, but that neither of these sets of sounds depends on the occurrence of sounds categorized by variable 3. A model with two pair-wise interactions such as [12,13] assumes that the occurrence of sounds categorized by variable 1 depends on the occurrence of sounds categorized by variables 2 or 3, but that the occurrence of sounds categorized by variables 2 and 3 are independent of one another, and so on.

The degrees of freedom of the models decrease progressively as interactions are added. For the 2x2x2 tables considered here, the degrees of freedom are: 4 for model [1,2,3], 3 for any of models [12], [13], or [23]; 2 for any of models [12,13], [12,23], or [13,23]; 1 for model [12,13,23]; and 0 for model [123].

As noted above, each model generates a set of expected frequencies for each cell in the table, and a G^2 statistic (twice the natural logarithm of the likelihood ratio statistic) is calculated from maximum-likelihood estimates of expected cell frequencies for each model. G^2, which is analogous and asymptotically equivalent to X^2 with very large samples like these, serves as a measure of the fit of expected to observed values for each model. As interactions are added, G^2 declines monotonically, indicating an increasingly better fit. (This monotonic decline is not a property shared by X^2 and is the reason for using G^2 instead.) Two criteria are employed in determining when to stop calculating further models: (1) when G^2 drops below the value indicating significance at $p < 0.05$ (for the number of degrees of freedom in that model) or (2) when the reduction in G^2 between any model with n interactions and any with $n+1$ interactions is not significant (for d.f. = 1). The model with the lowest G^2 beyond which no more significant reduction in G^2 occurs will be considered to fit the data best and used to interpret the effects of the variables on the co-occurrence dependencies, if any, among the sounds cross-classified by those variables.

7 References

Abbs, J.H., V.L. Gracco & K.J. Cole (1984). Control of multi-movement coordination: Sensory motor mechanisms in speech motor programming. *Journal of Motor Behavior* **16**. 195-231.

Al-Ani, S. (1970). *Arabic phonology*. The Hague: Mouton.

Alwan, A.A-H. (1986). *Acoustic and perceptual correlates of pharyngeal and uvular consonants*. MS thesis, MIT.

Bhat, D.N.S. (1973). Retroflexion: An areal feature. Stanford University, *Working Papers in Language Universals* **17**. 27-67.

Blumstein, S.E. (1986). On acoustic invariance in speech. In J.S. Perkell, J.S. & D.H. Klatt (eds.) *Invariance and variability in speech processes*. Hillsdale, NJ: Lawrence Erlbaum Assoc. 178-197.

Blumstein, S.E. & K.N. Stevens (1979). Acoustic invariance in speech production. *JASA* **66**. 1001-1017.

Broselow, E. & A. Niyondagara (1989). Feature geometry of Kirundi palatalization. *Studies in the Linguistic Sciences* **20**.

Browman, C.P. & L.M. Goldstein (1986). Towards an articulatory phonology. *Phonology Yearbook* **3**. 219-252.

Browman, C.P. & L.M. Goldstein (1988). Some notes on syllable structure in articulatory phonology. *Phonetica* **45**. 140-145.

Browman, C.P. & L.M. Goldstein (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M.E. Beckman (eds.) *Papers in laboratory phonology I: Between the grammar and physics of speech*. Cambridge, UK: Cambridge University Press. 341-376.

Browman, C.P. & Goldstein, L.M. (in press) Articulatory phonology: An overview. *Phonetica*.

Clements, G.N. (1976). Palatalization: Linking or assimilation. *CLS* **12**. 96-109.

Clements, G.N. (1985). The geometry of phonological features. *Phonology Yearbook* **2**. 225-252.

Clements, G.N. (1989). Place of articulation in consonants and vowels: A unified theory. *Working Papers of the Cornell Phonetics Laboratory* **5**. 77-123.

Collier, R. (1974). Laryngeal muscle activity, subglottal air pressure, and the control of pitch in speech. Haskins Laboratories, *Status Report on Speech Research* **SR-39/40**. 137-170.

Delattre, P.C. (1971). Pharyngeal features in the consonants of Arabic, German, French, Spanish and American English. *Phonetica* 23. 129-155.

Diehl, R.L. & J. Kingston (1991). Phonetic covariation as auditory enhancement: The case of the [+ voice]/[- voice] distinction. In O. Engstrand & C. Kylander (eds.) *Current Phonetic Research Paradigms: Implications for Speech Motor Control, PERILUS* 15. Stockholm: University of Stockholm. 139-143.

Diehl, R.L. & K.R. Kluender (1989). On the objects of speech perception. *Ecological Psychology* 1. 121-144.

Dorman, M.F., M. Studdert-Kennedy & L.J. Raphael (1977). Stop-consonant recognition: Release burst and formant transitions as functionally equivalent, context-dependent cues. *Perception & Psychophysics* 22. 109-122.

Dyhr, N.-J. (1991). The activity of the cricothyroid muscle and the intrinsic fundamental frequency of Danish vowels. *JPh* 47. 141-154.

Fienberg, S.E. (1980). *The analysis of cross-classified categorical data.* Cambridge, MA: MIT Press.

Fowler, C.A., P. Rubin, R. Remez & M.T. Turvey (1980). Implications for speech production of a general theory of action. In B. Butterworth (ed.) *Language Production I: Speech and Talk.* London, UK: Academic Press. 372-420.

Fowler, C.A. & M.T. Turvey (1980). Immediate compensation in bite-block speech. *Phonetica* 37. 306-326.

Ghazeli, S. (1977). *Back consonants and backing coarticulation in Arabic.* PhD dissertation, University of Texas, Austin.

Gnanadesikan, A.E. (1991). The feature geometry of coronal subplaces. MS. University of Massachusetts, Amherst.

Hayward, K.M. & R.J. Hayward (1989). 'Guttural': Arguments for a new distinctive feature. *Transactions of the Philological Society* 87. 179-193.

Honda, K. (1987). Relationship between pitch control and vowel articulation. In D.M. Bless & J.H. Abbs (eds.) *Vocal fold physiology: Contemporary research and clinical issues.* San Diego, CA: College-Hill Press. 286-297.

Honda, K. & Fujimura, O. (1991) Intrinsic vowel F_0 and phrase-final lowering: Phonological vs. biological explanations. In J. Gauffin & B. Hammerberg (eds.) *Vocal Fold Physiology: Acoustic, perceptual, and physiological aspects of voice mechanisms.* San Diego: Singular Publishing Group. 149-157.

Houde. R.A. (1968). *A study of tongue body motion during selected speech sounds, Monograph* 2. Santa Barbara, CA: Speech Communication Research Laboratory.

Hume, E. (1990). Front vowels, palatal consonants and the rule of umlaut in Korean. *NELS* **20**. 23-243.

Hume, E. (1992). *Front vowels, coronal consonants and their interaction in nonlinear phonology*. PhD dissertation. Cornell University.

Jakobson, R., G. Fant & M. Halle (1952). *Preliminaries to speech analysis*. Cambridge, MA: MIT Press.

Jackson, M.T.T. (1988). *Phonetic theory and cross-linguistic variation in vowel articulation*. University of California, Los Angeles, *Working Papers in Phonetics* **71**.

Jongman, A., S.E. Blumstein & A. Lahiri (1984). Invariant acoustic properties distinguishing alveolar and dental consonants. MS.

Keating, P.A. & A. Lahiri. (1990). Articulatory and acoustic differences between palatal and velar stops. *JASA* **88**. S80 (Abstract).

Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in consonants. *JASA* **73**. 322-335.

Kewley-Port, D., D. Pisoni & M. Studdert-Kennedy (1983). Perception of static and dynamic cues to place of articulation in initial stop consonants. *JASA* **73**. 1779-1793.

Kingston, J. (1991). Integrating articulations in the perception of vowel height. *Phonetica* **48**. 149-179.

Kingston, J. & A.H. Cohen (in press). Extending articulatory phonology: Comments on Browman and Goldstein's "Articulatory phonology: An overview". *Phonetica*.

Kingston, J. & R.L. Diehl (submitted). Phonetic knowledge and explanation: An essay on the nature of phonetic categories. *JPh*.

Klatt, D.H. & K.N. Stevens (1969). Pharyngeal consonants. MIT, Research Laboratory in Electronics, *Quarterly Progress Reports* **93**. 207-215.

Ladefoged, P. & P. Bhaskararao (1983). Non-quantal aspects of consonant production: A study of retroflex consonants. *JPh* **11**. 291-302.

Ladefoged, P., J. DeClerk, M. Lindau & G. Papcun (1972). An auditory-motor theory of speech production. University of California, Los Angeles, *Working Papers in Phonetics* **22**. 48-75.

Ladefoged, P. & I. Maddieson (1988). Phonological features for places of articulation. In L.M. Hyman & C.N. Li (eds.) *Language, speech and mind: Studies in honour of Victoria A. Fromkin*. London, UK: Routledge. 49-61.

Lahiri, A., L. Gewirth & S.E. Blumstein (1984). A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study. *JASA* **76**.391-404.

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. & I.G. Mattingly (1985). The motor theory of speech perception revised. *Cognition* **21**. 1-36.

Lindblom, B. (1983). Economy of speech gestures. In P. MacNeilage (ed.) *The production of speech*. New York, NY: Springer Verlag. 217-245.

Lindblom, B. (1986). Phonetic universals in vowel systems. In J.J. Ohala & J.J. Jaeger (eds.) *Experimental phonology*. Orlando, FL: Academic Press. 13-44.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W.J. Hardcastle & A. Marchal (eds.) *Speech production and speech modelling*. Dordrecht: Kluwer Academic Publishers. 403-439.

Lindblom, B., J.F. Lubker & T. Gay (1979). Formant frequencies of some fixed mandible vowels and a model of predictive simulation. *JPh* 7. 147-162.

Lindblom, B. & I. Maddieson (1988). Phonetic universals in consonant systems. In L.M. Hyman & C.N. Li (eds.) *Language, speech and mind: Studies in honour of Victoria A. Fromkin*. London, UK: Routledge. 62-78.

Lindblom, B. & J. Sundberg (1971). Acoustical consequences of lip, tongue, jaw and larynx movement. *JASA* 50. 1166-1179.

Lubker, J.F. (1968). An electromyographic-cinefluorographic investigation of velar function during normal speech production. *Cleft Palate Journal* 5. 1-18.

MacNeilage, P. (1970). Motor control of serial ordering of speech. *Psychological Review* 77. 182-196.

McCarthy, J.J. (1988). Feature geometry and dependency: A review. *Phonetica* 45. 84-108.

McCarthy, J.J. (1989). Guttural phonology. MS. University of Massachusetts, Amherst.

McCarthy, J.J. (in press). The phonetics and phonology of Semitic pharyngeals. In P. Keating (ed.) *Papers in Laboratory Phonology* III. Cambridge, UK: Cambridge University Press.

Maddieson, I. (1983). The analysis of complex phonetic elements in Bura and the syllable. *Studies in African Linguistics* 14. 285-310.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge, UK: Cambridge University Press.

Mann, V. & B.H. Repp (1980). Influence of vocalic context on perception of the [ʃ]-[s] distinction. *Perception & Psychophysics* 28. 213-228.

Moll, K.L. (1962). Velopharyngeal closure in vowels. *Journal of Speech and Hearing Research* 5. 30-37.

Moll, K.L. & T.H. Shriner (1967). Preliminary investigation of a new concept of velar activity during speech. *Cleft Palate Journal* 4. 58-69.

Nearey, T.M. (1990). The segment as a unit of speech perception. *JPh* 18. 347-373.

Ohala, J.J. (1970). *Aspects of the control and production of speech.* University of California, Los Angeles, *Working Papers in Phonetics* **15**.

Ohala, J.J. (1978). Southern Bantu vs. the world: The case of palatalization of labials. *BLS* **4**. 370-386.

Ohala, J.J. (1979). The contribution of acoustic phonetics to phonology. In B. Lindblom & S. Öhman (eds.) *Frontiers of speech communication research*. London, UK: Academic Press. 355-363.

Ohala, J.J. (1983). The origin of sound patterns in vocal tract constraints. In P. MacNeilage (ed.) *The production of speech*. New York, NY: Springer Verlag. 189-216.

Ohala, J.J. (1985). Around *flat*. In V. Fromkin (ed.) *Phonetic linguistics: Essays in honor of Peter Ladefoged*. Orlando, FL: Academic Press. 223-241.

Ohala, J.J. & B.M. Eukel (1987). Explaining the intrinsic pitch of vowels. In R. Channon & L. Shockey (eds.) *In honour of Ilse Lehiste*. Dordrecht: Foris. 207-215.

Ohala, J.J. & J. Lorentz (1977). The story of [w]: An exercise in the phonetic explanation for sound patterns. *BLS* **3**. 577-599.

Padgett, J. (1991). *Stricture in feature geometry*. PhD dissertation. University of Massachusetts, Amherst.

Perkell, J.S. (1969). *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. Cambridge, MA: MIT Press.

Peterson, G.E. & H.L. Barney (1952). Control methods used in a study of vowels. *JASA* **24**. 175-184.

Recasens, D. (1983). Place cues for nasal consonants with special reference in Catalan. *JASA* **73**. 1346-1353.

Repp, B.H. (1982). Phonetic trading relations and context effects: New evidence for a phonetic mode of processing. *Psychological Bulletin* **92**. 81-110.

Repp, B.H., A.M. Liberman, T. Eccardt & D. Pesetsky (1978). Perceptual integration of acoustic cues for stop, fricative and affricate manner. *Journal of Experimental Psychology: Human Perception and Performance* **4**. 621-637.

Riordan, C. (1977). Control of vocal tract length in speech. *JASA* **62**. 998-1002.

Sagey, E. (1986). *The representation of features and relations in nonlinear phonology*. PhD dissertation. MIT.

Shadle, C.H. (1985). *The acoustics of fricative consonants*. PhD dissertation, MIT (published as Research Laboratory of Electronics, *Technical Report* **506**.)

Stevens, K.N. (1972). The quantal nature of speech: Evidence form articulatory-acoustic data. In David & P. Denes (eds.) *Human communication: A unified view*. New York, NY: McGraw-Hill. 51-66.

Stevens, K.N. (1989). On the quantal nature of speech. *JPh* 17. 3-45.

Stevens, K.N. & S.E. Blumstein (1975). Quantal aspects of consonant production and perception: A study of retroflex stop consonants. *JPh* 3. 215-233.

Stevens, K.N. & S.E. Blumstein (1978). Invariant cues for place of articulation in stop consonants. *JASA* 64. 1358-1368.

Sussman, H.M., H.A. McCaffrey & S.A. Matthews (1991). An investigation of locus equations for stop place categorization. *JASA* 90. 1309-1325.

Sussman, H.M., F. Ahmed & K. Hoemeke (1992). A cross-linguistic validation of locus equations: Are there "phonetic hot spots" in CV acoustic space? *JASA* 92. S2413. (Abstract)

Syrdal, A. & H. Gopal (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America* 79. 1086-1100.

Tuller, B. & H.L. Fitch (1980). Preseveration of vocal tract length: A negative finding. *JASA* 67. 1068-1071.

Vilkman, E., O. Aaltonen, I. Raimo, P. Arajärvi & H. Oksanen (1989). Articulatory hyoid-laryngeal changes vs. cricothyroid activity in the control of intrinsic F_0 of vowels. *JPh* 17. 193-203.

Whalen, D.H. (1981). Effects of vocalic formant transitions and vowel quality on the English [s]-[ʃ] boundary. *JASA* 69. 275-282.

Whalen, D.H. (1989). Vowel and consonant judgments are not independent when cued by the same information. *Perception & Psychophysics* 46. 284-292.

Whalen, D.H. (1992). Perception of overlapping segments: Thoughts on Nearey's model. *JPh* 20. 493-496.

Wood, S. (1982). *X-ray and model studies of vowel articulation*. Lund University, Department of Linguistics, *Working Papers* 23.

Zakia, R.A. (1991). *The role of transition duration in the perception of speech and nonspeech*. PhD dissertation. Cornell University.

Zvelebil, K. (1970). *Comparative Dravidian Phonology*. The Hague: Mouton.

Notes

* The preparation of this paper was aided substantially by discussions with John McCarthy and Elisabeth Selkirk and by the comments of the audience at the Language and Cognition Colloquium, Northeastern University, 27 May 1992. The author remains entirely responsible for its contents, however.

1. As both the appearance of unmarked sounds in neutralization contexts and their absence from phonotactic constraints in a language often correlates with their higher crosslinguistic frequency, the relative markedness of a sound can often be determined by multiple, independent tests.

2. Heavy nasalization simply exacerbates these effects with high vowels, but actually reverses the effect with low vowels, by raising the amplitude of the nasal pole relative to the oral pole above it. This shifts the center of gravity downward in the low frequency range and thus makes a low vowel sound higher.

3. The redundancy of lip rounding with tongue positions that are phonologically [+back] is actually more complex that this, since low, back vowels are frequently unrounded, and rounding tends to increase with height among the (non-low) back, rounded vowels (Linker, 1982). Both the categorical and the continuous interactions of lip rounding with vowel height may also reflect the speaker's deliberate manipulation of articulations to produce desirable perceptual interactions. First, low vowels may be left unrounded so that lip rounding doesn't lower the high F_1 produced by the pharyngeal constriction. Second, the higher a back vowel is the further front along the palate it's actually articulated (Perkell, 1969; Wood, 1982; Jackson, 1988), and thus the shorter the cavity in front of the constriction. More lip rounding will therefore be required to lengthen this cavity and thereby lower F_2 enough in the higher back vowels.

4. Ohala (1983) presents a similar case in arguing that because of the low pressure buildup behind a bilabial stop closure, the burst of noise at the release of a /p/ will be quite weak and that sound may be easily confused with an /f/ or /ɸ/. More formally, the [continuant] contrast is tenuous for [labial, -voice, -sonorant] sounds, and will collapse if the integration of the consonant's manner with its place prevents a lowering of the [-continuant] threshold.

5. Another demonstration that articulations are motor equivalent involved loading the lower lip when the speaker is trying to close the lips (Abbs, Gracco & Cole, 1984). Speakers immediately compensated with more lowering of the upper lip, as well as applying more force to raise the lower lip against the load. A third case showing compensation by additional larynx lowering for a impediment to lip protrusion in rounded vowels is discussed below.

6. The task dynamics also transforms the instantaneous onsets and offsets of gestures into gradual, continuous articulatory movement.

7. Tuller and Fitch (1980) failed to obtain additional larynx lowering from American English speakers who they asked not to protrude their lips when producing rounded vowels. There are a number of differences between their study and Riordan's that could explain their failure to replicate. First, unlike Riordan's subjects, their speakers' lips were not actually constrained from protruding. Second, Riordan examined the larynx lowering in speakers of French and Mandarin who may produce more lip rounding in their rounded vowels than English speakers do, and may thus need to lower the larynx more when lip protrusion is constrained. The larynx lowering reported by Lindblom and Sundberg (1971) as an accompaniment to lip rounding is apparently based on observations of Swedish speakers, who also are described as rounding their lips more than English speakers.

8. In the following discussion, the symbols for voiceless, unaspirated stops are used to refer to all stops at a particular place of articulation, and laryngeal articulations were ignored entirely in determining the places at which a language has stops. This means that if a language had a stop with any laryngeal articulation at a particular place, it was counted as having a stop at that place even if it has no stops with the same laryngeal articulation at other places. Three [+anterior, coronal] places are potentially distinguished in UPSID: sounds which are explicitly described in the sources as "dental" or "alveolar" and sounds which are "uncertain dental/alveolar." This last category is used when the source is not specific as to whether the stop is dental or alveolar. All cases of this type will be treated as alveolar here, i.e. a sound will only be treated as dental if UPSID (and its sources) are explicit that it's actually dental. This approach, arbitrary as it may be, causes few problems since if a language has just one [+anterior, coronal] stop, it doesn't matter for my purposes whether it's genuinely alveolar or actually dental, and if it has more than just one, UPSID (and its sources) reliably distinguish where necessary between dental and alveolar stops in nearly all cases.

9. Hawaiian has no dental or alveolar stop; Wichita, Aleut, and Hupa have no bilabial stop; and Hupa and Kirghiz have no velar stop, although both have a uvular stop.

10. This exclusion reflects the fact that this paper focuses generally on the question of what articulations are common, not those which are merely possible. Most discussions of phonological representations of place contrasts have confounded frequency with possibility and thus do not provide a comprehensive assessment of the relative markedness of the various places. Furthermore, Maddieson (1983) has recently presented evidence that the apparent labio-coronals of Bura may actually be clusters.

11. In the following discussion, "high frequency" refers to frequency region near the neutral value of F_4, while "low frequency" refers to the frequency region near the neutral frequency of F_2.

12. This is true both of spectra of the noise in the fricative /χ/ and of spectra of the burst at the release of the stop /q/. Unfortunately, Alwan doesn't present spectral data on velar consonants and following vowels produced by these speakers, so it's not possible to assess whether front vowels are backed after a uvular consonant compared to after a velar. There are, however, at least two descriptions (Ghazeli, 1977; Al-Ani, 1970) that show that the three uvular consonants /q,ʁ,χ/ in Arabic differ in how far back the articulation is made: with /q/ being furthest back, /χ/ furthest front, and /ʁ/ in between (Ghazeli's data is for a Tunisian speaker and Al-Ani's for an Iraqi speaker). These differences predict that F_2 at a following vowel onset should be lowest after /q/, highest after /χ/, and intermediate after /ʁ/. These predictions are only partially borne out by the F_2 onset frequencies reported by Alwan for her four speakers: F_2 was consistently lower after /ʁ/ than after /χ/. But after /q/, F_2 could be higher or lower than after either or both of the other two uvular consonants, suggesting either that the uvular stop is not consistently further back than the fricative for these speakers or that the manner contrast also affects F_2. Theoretical modeling (Klatt & Stevens, 1969; Alwan, 1986) suggests, however, that the closer constriction of the stop should lower F_2 more than the more open constrictions of the fricatives, especially the voiced uvular fricative /ʁ/ which is often open enough to be (nearly) an approximant rather than a fricative, so it's unlikely that /q/'s effect on F_2 is a product of its manner.

13. A hypothesis quite similar to the one promoted here, that a lack of change in the energy distribution in the spectrum between the stop release and vowel onset is a positive property for identifying velar consonants, has been proposed by Kewley-Port (1983; Kewley-Port, Pisoni & Studdert-Kennedy, 1983).

14. Stevens and Blumstein (1978) show that the three major places are more reliably distinguished by differences in formant transitions alone rather than by differences in the short-time spectra of stop releases alone (differences in both formant transitions and release short-time spectra, of course, distinguish the three places better than either kind of difference alone). Even more dramatic evidence that spectral change is essential differentiator of places can be found in Sussman's recent work on locus equations (Sussman, McCaffrey & Matthews, 1991; Sussman, Ahmed & Hoemeke, 1992). Locus equations represent the linear relationship between a formant's frequency at the onset (or offset) of its transition from a consonant into a vowel and its steady-state frequency in the vowel. Sussman et al. (1991) show that the *y*-intercept and

slope of lines drawn between F_2's onset and steady-state frequencies reliably differentiate stops produced at three major places of articulation before 10 different vowels by ten female and ten male English speakers. The locus equation for the [labial] stop /b/ has a lower *y*-intercept and steeper slope than do the locus equations for the [coronal] stop /d/ or the [dorsal] stop /g/. Two locus equations are needed for the [dorsal] stop /g/: the one for /g/'s before back vowels has a lower *y*-intercept and steeper slope than that for /g/'s before front vowels. Finally, the locus equation for the [coronal] stop /d/ has the shallowest slope and a *y*- intercept between that for /g/ before back and front vowels. Sussman et al. (1992) show that plain and uvularized [coronal] stops in Cairene Arabic and dental and retroflex stops in Urdu are also reliably distinguished by their locus equations. This success suggests that there are after all acoustic invariants for place, though they are not necessarily quantal in nature (cf. Stevens, 1972, 1989). These results also suggest that the identification of a stop's place does not require that the listener match an acoustic pattern to the articulation that produced, contrary to the claims of the motor theory of speech perception (Liberman, Cooper, Shankweiler & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985). All these results suggest, too, that dynamic spectral properties are superior to static ones in distinguishing stops for place.

15. However, the [coronal] stop in fully 124 of the languages with just one stop made with this articulator is classified as uncertain dental/alveolar, meaning that UPSID's sources for those languages didn't specify whether the stop was dental or alveolar. Since this means that nearly half of the languages with just one [coronal] stop are left out of this comparison, the genuine frequency of dentals vs. alveolars as the sole [coronal] stop is uncertain.

16. This account supports Gnanadesikan's (1991) analysis of retroflex stops as a [coronal] articulation with a superimposed [+back] specification. This analysis gains further support from the fact that retroflexes develop historically from alveolars in the environment of a (preceding) back vowel in both Dravidian and Australian languages (see Zvelebil, 1970; Bhat, 1973; Kingston & Cohen, in press, for further discussion and exemplification).

17. This discussion has not addressed contrasts between consonants made with laminal vs. apical contact, i.e. those represented by the feature [distributed], which are typically treated as a variety of place contrasts. For example, dental consonants are often (though not always (Ladefoged & Maddieson, 1988)) laminal in contrast to apical alveolars, and retroflexes are generally apical in contrast to laminal palato-alveolars. The reason for ignoring [distributed] contrasts is simply that this paper's focus is on place contrasts *sensu strictu*, i.e. on articulatory differences in the location of a constriction,

which affect what were called in the Trubetskoian-Jakobsonian tradition the "tonality" of the consonant. Differences of the sort represented by such features as [distributed] are instead differences in the manner in which a constriction is made at a location. That [distributed] is a dependent of [coronal] (see Gnanadesikan (1991) for convincing arguments) doesn't require that it be treated as a place feature if Padgett (1991)'s arguments for treating manner features generally as dependents of place nodes are accepted.

18. The absence of languages with true palatals and either palatalized alveolar or palatalized velars is not a product of the overall rarity of these sets of sounds in UPSID; they are each frequent enough that at least 6 languages are expected to have a true palatal and one of the palatalized sounds. A comparison of the observed vs. expected distribution of these sets of sounds by a X^2 test shows that a significant difference ($p < 0.05$) between them.

19. The numbering of the variables in these analyses will always follow the pattern in table 3 below, in which the contrast referred to as variable [3] is at the top, [2] below it, and [1] on the side; the numbering of variables is arbitrary and no priority is implied for higher or lower numbered variables.

20. If this hypothesis is rejected, then the mutual exclusion of these two classes of sounds undermines the independence of [coronal] and [dorsal] articulators, and of their subordinate features, since a particular specialization of a [coronal] articulation in a language, i.e. a palato-alveolar, blocks the co-occurrence of a specialization of a [dorsal] articulation, i.e. a palatal, and vice versa. The representational difficulty here is reminiscent of that presented by the natural class of consonants described as "gutturals" by McCarthy (1989, in press, see also Hayward & Hayward, 1989), which include sounds articulated from the glottis to the uvula. Since these sounds are made with a variety of articulators, true and false vocal folds and the tongue root and dorsum, McCarthy suggests that this natural class of sounds is unified by sharing the same (rather broad) place of articulation (which he calls "pharyngeal") rather than because they are all made with the same articulator. The same move could be made here, uniting palato-alveolars and palatals as sharing a post-alveolar place of articulation.

21. Articulatory cost is conceived more abstractly here than in Lindblom (1983), where articulatory cost increases concretely with the degree that an articulator is displaced from its current (or a rest) position. Here, articulatory cost instead increases when a new articulation is employed, but not when one that is already used is simply used in a new way.

22. As indicated in this table, /f,ɸ/ were pooled in the counts of [labial] fricatives, as were /ʃ,θ,s̪,ɬ/ in the counts of [coronal] fricatives, and /ç,x,χ/ in the counts of [dorsal] fricatives (laryngeal articulations were again ignored

in the counts). One member of each of these classes occurs much more frequently than any of the others, i.e. /f/, /ʃ/, and /x/, and similar results are obtained when these representatives of fricatives made with each articulator are considered alone. The symbols of these sounds are also used as shorthand to represent the entire class in the discussion below.