Acoustic and Perceptual Evidence for Universal Phonological Features

Kenneth N. Stevens
Research Laboratory of Electronics and Department of Electrical Engineering and Computer Science,
Massachusetts Institute of Technology, Cambridge MA 02139-4307, USA
E-mail: stevens@speech.mit.edu

ABSTRACT

The human speech production apparatus consists of a number of articulators or structures that have the capability of producing somewhat continuous movements within particular ranges. However, with proper adjustment of postures of the respiratory system and of certain other articulators, the properties of the sound that is generated by these movements and the human responses to this sound tend to fall into discrete categories. These categories form the bases for the distinctive features that potentially define the phonological contrasts in any language. This quantal aspect of articulatory/acoustic/perceptual relations arise from the properties of coupled resonators in the vocal tract, the nature of the sound sources in the vocal tract, the anatomy and physiology of vocal-tract structures, and auditory responses to sounds with spectral prominences and to the temporal properties of these sounds. In a given language, the perceptual saliency of a particular contrast may be enhanced by introducing articulatory gestures that provide acoustic cues for the contrast over and above the cues resulting from the defining gestures. Examples of the language-independent defining properties for a variety of features are given, as well as examples of language-dependent enhancing gestures and acoustic cues.

1. BACKGROUND

A basic component of any language is a lexicon of words. These words, in turn, can be represented as sequences of underlying segments, and the segments can be described by bundles of discrete distinctive features. A change in the value of a distinctive feature in a segment of a word has the potential of creating a different word. Examples of minimal pairs of words that result from changing one feature of a segment are cat/pat, cap/cab, and dim/deem. The feature in each case is distinctive in the sense that it defines a contrast in the language [1],[2]. These distinctive features are assumed to be binary; that is, a feature can have a value of "+" or "-". Many linguists accept the premise that there is a universal inventory of these distinctive features, and that any one language utilizes a subset of these features to form the contrasts between words.

What is the basis for this proposed universal inventory of distinctive features? This paper will review the evidence for features that emerges from an examination of the relation between the articulatory gestures that result when a word with a particular feature in one of its feature bundles is produced, the acoustic consequence of these articulatory gestures, and the auditory response to the acoustic patterns.

2. SOME EXAMPLES

We consider in some detail the defining acoustic/perceptual and articulation bases for four of the distinctive features. The discussion of these four features will illustrate the arguments used in developing similar arguments for many if not all of the other features.

2.1 THE FEATURE [BACK] FOR VOWELS

The articulatory correlate for vowels that are [+back] is a tongue-body position that is backed, so that a narrowing of the airway occurs either in the pharyngeal region or in the posterior part of the oral region. The acoustic correlate for this vocal-tract shape is a relatively low second formant frequency F2. It is usually assumed that during the production of vowels, the impedance of the glottal source is large compared with the impedance looking up into the vocal tract, so that there is little acoustic interaction between the supraglottal cavities and the subglottal system. While this assumption is usually valid over most frequencies, it is less valid when the supraglottal resonances or formants are at the frequencies of the resonances of the subglottal system, particularly the second subglottal resonance. For an adult male or female, this resonance (which we will call fsub2) is roughly in the range 1500 to 1700 Hz [3],[4]. When there is a natural frequency of the supraglottal airway in this frequency range, an interaction between the sub- and supraglottal resonances can usually be observed. The amount of interaction depends on the average glottal opening during phonation, and is greater for some speakers than for others. If we call F2 the second natural frequency of the vocal tract in the absence of the influence of the subglottal system, then when F2 is close to fsub2, the spectrum of the vowel in the vicinity of F2 is modified. For F2 well below or well above fsub2, the frequency and amplitude of the spectrum prominence for F2 is not significantly modified by fsub2. However, when F2 is close to fsub2, there is a spectrum prominence slightly below F2 or slightly above F2, but not exactly at F2. Thus if the vocal tract were to be positioned such that F2 is very close to fsub2, the frequency of the

 ISBN 1-876346-48-5

spectrum prominence in the vicinity of F2 is somewhat unstable. Some evidence for this instability can be seen in diphthongs like /αi/, in which F2 passes through fsub2. For many if not most speakers, a discontinuity in the frequency and amplitude of the F2 spectrum prominence can be seen at this point in the F2 movement [5]. An example for a female speaker is given in Fig. 1, which shows measured formant movements (obtained from linear prediction analysis) during production of the word bide. Discontinuities are seen in the F2 track at about 1800 Hz, and in the F1 track at about 700 Hz, which is near the lowest natural frequency of the subglottal system. There is, in a sense, a natural boundary that defines the opposition between [+back] and [-back] vowels.

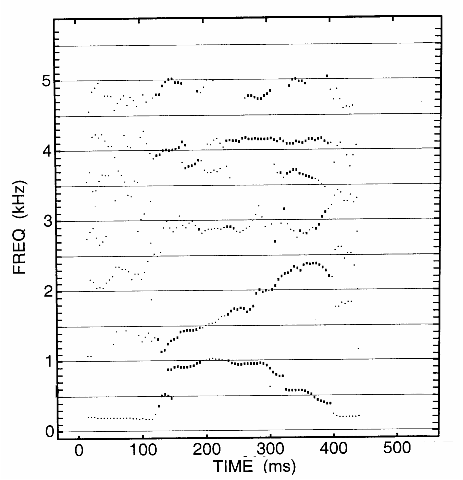

Figure 1: Trajectories produced by an LPC-based formant tracker for the word bide produced by a female speaker. There are discontinuous jumps as the formants pass through tracheal resonances for this speaker.

2.2 THE FEATURE [CONTINUANT]

The feature [continuant] is distinctive for obstruent consonants, and distinguishes between stop consonants ([-continuant]) and fricative consonants ([+continuant]). For [-continuant]) segments, an articulator makes a complete closure in the oral cavity. In common with all obstruent consonants there is a buildup of pressure above the glottis, causing the transglottal pressure to decrease and resulting in either cessation of glottal vibration or greatly weakened glottal vibration. A [-continuant] consonant is distinguished from a [+continuant] consonant by the lack of acoustic energy in the frequency range above the first one or two harmonics during the constricted interval. At the release of a stop consonant there is usually a brief interval of frication noise produced by the rapid flow of air through the constriction formed by the moving articulator. For a [+continuant] consonant, the cross-sectional area of the constriction formed in the oral cavity is in the range that generates rapid airflow in the constriction, giving rise to the generation of turbulence noise in the vicinity of the constriction or at an obstacle downstream from the constriction. Analysis shows that, for a given glottal opening, there is a range of constriction areas for which the amplitude of the turbulence noise source is relatively insensitive to the cross-sectional area of the oral constriction [6]. A theory-derived relation between the amplitude of turbulence noise generated near a consonantal constriction and the cross-sectional area of the constriction (assuming, the glottal area is modified by the intraoral pressure) is shown in Fig. 2. For smaller constriction sizes the airflow is smaller, and therefore generates less noise. For larger constriction areas, the airflow is limited by the size of the glottal opening, and the velocity of the flow at the constriction decreases as its area increases. Again, then, we have an example of the prototypical articulatory/acoustic relation that underlies a feature, where there is stable region and, at the edges of this region, a condition where the acoustic and perceptual result is sensitive to small changes in articulation.

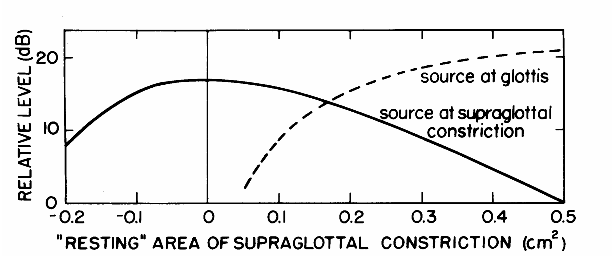

Figure 2: Calculated relative levels of noise sources at glottis and at supraglottal constrictions as a function of the supraglottal constriction area A_c that would be obtained if there were no forces on the constriction due to the intraoral pressure. For a fixed glottal area, there is a range of A_c in which the noise amplitude is relatively insensitive to A_c.

2.3 THE FEATURE [NASAL]

At first glance, the defining attributes for the feature [nasal] seem straightforward, at least in its articulatory implementation. While the default configuration for the soft palate is a raised position, with a closed velopharyngeal opening, this articulator is lowered for segments that are [+nasal], creating a velopharyngeal opening. In the case of vowels, a sufficient degree of opening of the velopharyngeal port has several acoustic consequences, a principal one being a decreased prominence of the first-formant spectrum peak, and another, related, attribute is the appearance of additional spectrum prominences in the vicinity of the first formant [7],[8].

Some reflection shows, however, that there is, in a sense, an optimal range of cross-sectional area of the velopharyngeal opening when the nasal-nonnasal distinction applies to vowels, and possibly a different range of this opening for consonants that are distinctively nasal. In the case of vowels, the optimal influence of the velopharyngeal opening on the spectrum will occur when the impedance looking into the nasal cavity is about the same magnitude as the impedance looking back into the vocal tract. This means that the area of the velopharyngeal opening should be roughly equal to the minimum cross-sectional area of the vocal tract during a vowel, on average. This area is probably about 1 cm^2. For nasal consonants, a complete closure is formed by one of the oral articulators, and the sound is radiated from the nostrils. In the normal case of a voiced nasal consonant, there is very little pressure increase behind the oral constriction, so that the cross-sectional area of the velopharyngeal part must be sufficiently large to prevent this pressure increase. This minimum cross-sectional area should be about 0.3 cm^2. On the other hand, too large a velopharyngeal opening could result in a decreased influence of the oral cavity on the spectrum of the sound, and hence a decreased ability to produce perceptually distinctive sounds for different places of articulation. Research with a speech synthesizer that simulates the acoustic effect of the nasal cavity for consonants suggests an optimum area of the velopharyngeal port to be in the range 0.3 to 0.6 cm^2. Thus the range of velopharyngeal openings appropriate for nasal consonants tends to be smaller than that for distinctively nasal vowels.

2.4 THE FEATURE [ANTERIOR] FOR OBSTRUENT CONSONANTS

The various place of articulation features for consonants include features that specify which of three articulators is the primary articulator that forms the constriction, and hence is responsible for creating the acoustic discontinuity or abruptness that is characteristic of consonants. Within the class of consonants for which the tongue blade is the primary articulator, distinctions are made in some languages by modifying the position of the tongue-blade constriction along the hard palate and by manipulating the shape of the tongue blade. We examine here the articulatory and acoustic consequences of the tongue blade positioning. In the case of obstruent consonants, we derive the properties of the radiated frication noise when there is rapid airflow at the constriction leading to acoustic excitation of certain natural frequencies of the vocal tract, particularly the natural frequencies of the cavity anterior to the constriction.

The configuration of the vocal tract when a constriction is made with the tongue blade can be modeled by a tube that is closed at one end and open at the other with a narrowing in the region from 1-5 cm from the open end, i.e., with a front-cavity length ℓ_f of 1-5 cm. To illustrate the acoustic behavior of such a tube, we assume the length of the constriction to be constant, and the length ℓ_b of the back cavity is adjusted so that the total length is constant for different values of ℓ_f. If the cross-sectional area of the constriction is very small, then there is negligible acoustic interaction between the two parts of the tube, and the natural frequencies of the combination are simply of the individual parts. Thus the back cavity contributes the frequencies $c/(2\ell_b)$, c/ℓ_b, $3c/(2\ell_b)$, etc., and the front cavity contributes $c/(4\ell_f)$, where c=velocity of sound. For ℓ_f in the range 1-5 cm, these calculated natural frequencies are shown by the lines in Fig. 3.

We now assume the cross-sectional area of the constriction to be somewhat larger, say about 0.1 cm^2, as might occur for a fricative consonant. The natural frequencies given by the dashed lines in Fig. 3 are now modified, and take on values given by the solid lines. The frequencies tend to be shifted up slightly when the front and back cavity resonances are well separated, but when the two resonances are close together, i.e., near the points of intersection in the figure, there is a substantial modification. If one were to trace the movements of the high-frequency spectrum prominence for a fricative as the front cavity length increases, there would be a discontinuous shift in the frequencies of the prominence, shown by the connecting lines. As ℓ_f increases, the frequency of the spectrum prominence for this model would be in the range 6 to 7 kHz for ℓ_f around 1.5 cm (corresponding to the 6^{th} formant); 4.5-5.5 kHz for ℓ_f around 1.7 cm (5th formant); 3.2 to 4 kHz for ℓ_f around 2.5 cm (4^{th} formant); and 2-2.9 kHz for ℓ_c around 3.5 cm (3^{rd} formant). Values of ℓ_f near the boundaries of these regions would lead to changes in the frequency of the spectrum prominence for small changes in ℓ_f, as shown by the thick lines.

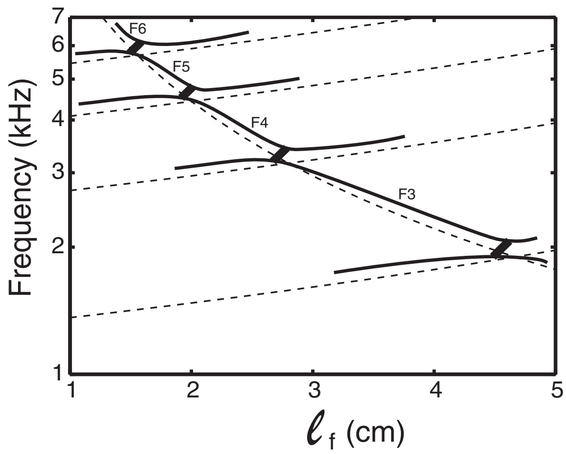

Figure 3: Calculated natural frequencies for a model of a vocal tract with a narrow constriction as the length ℓ_f of the front cavity changes from 1 to 5 cm. The dashed lines give the frequencies of the back- and front-cavity resonances assuming no acoustic coupling between these cavities. The solid lines are the resonances if coupling is taken into account. See text.

3. PRINCIPLES UNDERLYING QUANTAL RELATIONS

The examples given in Section 2, and many other examples, suggest that there are natural boundaries in acoustic/perceptual space when ranges of values for particular types of articulatory gestures are produced [9],[10]. We explore in this section the general nature of articulatory/acoustic/perceptual relations for the vocal tract, the acoustic principles that give rise to these apparent boundaries, and the particular anatomical and physiological aspects of the human vocal tract that appear to favor the emergence of the boundaries. In this paper we are concerned primarily with articulatory-acoustic relations, although the auditory system certainly also plays a central role in shaping these natural boundaries.

 ISBN 1-876346-48-5 © 2003 UAB

3.1 ARTICULATORY/ACOUSTIC/PERCEPTUAL RELATIONS ARE NONMONOTONIC

In all of these examples, the relation between an acoustic parameter and a parameter that specifies an articulatory gesture shows a somewhat nonmonotonic or discontinuous form. Over one set of values for the dimension describing the articulatory manipulation the resulting acoustic attribute is relatively stable. As the articulatory parameter extends beyond this range of values, there is a rather abrupt change in the acoustic parameter into a new value --- a value which gives rise to a significantly different auditory response. The articulatory-acoustic relation can be schematized in the manner shown in Fig. 4. Two aspects of this relation are significant: (1) Over particular ranges of the articulatory parameter (labeled I and III in Fig. 4) there is not much variation in the resulting acoustic parameter. Put another way, some variability in the articulatory parameter can be tolerated without significantly changing the acoustic result. (2) There is a boundary in the acoustic-articulatory relation where the acoustic parameter changes rather abruptly. Setting an articulatory parameter in region II can lead to an unstable condition where small changes in articulation lead to relatively large changes in the acoustics, and presumably, the perception. For each of the examples given in Section 2, relations of this kind are obtained, with different acoustic and articulatory parameters.

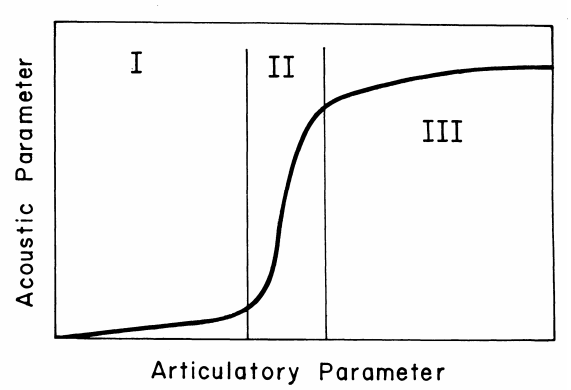

Figure 4: Schematic representation of the relation between an acoustic and an articulatory parameter as the articulatory parameter is manipulated through a range of values.

3.2 ACOUSTIC PRINCIPLES RESPONSIBLE FOR QUANTAL RELATIONS

All of these “quantal” relations between acoustic and articulatory parameters arise from three acoustic principles: theories of coupled acoustic resonators, the effects of vocal-tract walls and of acoustic energy losses on prominences of the system, and principles of sound source generation due to airflow at constrictions in the vocal tract.

In the examples given in Section 2, the coupling between resonators plays a role in establishing the boundaries (as in region II of Fig. 4) for several articulatory-acoustic relations. These include coupling between the oral cavity and the nasal cavity, coupling between the subglottal system and the vocal tract above the glottis, and coupling between portions of the vocal tract that are separated by a narrow constriction.

The second acoustic principle is concerned with acoustic consequences of energy losses and the acoustic effects of yielding walls in the vocal tract. If the vocal tract is configured such that there are increased acoustic losses, either at the glottis and in the subglottal system, in the nasal cavity, or at a constriction in the vocal tract, then the spectrum prominence for one or more formants will be decreased. For example, there are significant acoustic energy losses at the walls of the nasal cavity for low and mid frequencies, leading to an increased bandwidth of the first formant and a reduced prominence of F1 --- a principal acoustic and perceptual cue for nasalization of vowels. Increased bandwidth of some formants also occurs for tense high vowels. This increased bandwidth for the front vowel usually occurs for F3 due to radiation resistance. For the back vowel /u/ the bandwidth of F2 appears to be caused by acoustic resistance created by the tongue body and lip constrictions. The consequence is a tendency for two nearby formants to create a single spectrum prominence --- F3 and F4 for the high front vowel, and F1 and F2 for the high back vowel. The acoustic effect of yielding walls on F1 for high vowels also contributes to the stability of F1 for such vowels. The acoustic mass of the vocal-tract walls makes F1 less sensitive to the area of the tongue-body constriction than it would be if the vocal-tract walls were hard. And when the glottis is somewhat open, as it might be in a vowel under certain conditions such as in the context of /h/, there is an increased bandwidth (and decreased prominence) of the first formant. This reduced prominence of the F1 spectrum peak contributes to the identification of features related to glottal spreading.

The third acoustic principle is related to the mechanisms by which sound sources are generated during speech production. The glottal source for phonation can be influenced by several factors, including the transglottal pressure, the vocal-fold stiffness, and the degree of glottal abduction or adduction. There are ranges of variation for each of these parameters that facilitate glottal vibration, and outside of these ranges vibration fails to be initiated. The intensity of turbulence noise in the vicinity of a constriction, either at the glottis or in the supraglottal region, also shows quantal attributes, as observed above in Fig. 3.

3.3 ANATOMY AND PHYSIOLOGY FAVORING QUANTAL RELATIONS

There are several aspects of the anatomy and physiology of the speech production system that make it possible for articulatory-acoustic relations of the type schematized in Fig. 4 to occur. For example, the bend at about the midpoint in the vocal tract, from the pharyngeal to the oral region, with the soft palate located at the bend, provides several favorable conditions. As a consequence of this geometry, front-back tongue body movements are well correlated with F2, and this strengthens the emergence of the feature [back] for vowels. It also helps to define place-of-articulation features for nasal consonants, which depend on the distance between the velopharyngeal opening and the location of the consonant closure. The particular characteristics of the yielding walls of the vocal tract make it possible to produce stable constrictions for fricative consonants and to produce cue-generating noise bursts at the release of stop consonants [11]. The fact that the first and second resonances of the subglottal system fall within the range of the first and second natural frequencies of the vocal tract is, of course, related to the dimensions of the structures, and, as noted above, helps to provide a basis for some of the place features for vowels. And the particular anatomy of the nasal cavity makes it possible to exploit the additional acoustic losses in this cavity in providing cues for the feature [nasal], particularly for vowels. If one takes the view that the existence of an inventory of discrete phonological categories is one of the requirements for language, then it might be argued that the anatomical features of the human vocal tract favor the emergence of a number of these categories.

4. VARIABILITY DUE TO ENHANCEMENT AND OVERLAP

In Section 2 we gave examples of articulatory-acoustic (and hence articulatory-perceptual) relations that exhibit quantal characteristics. We propose that quantal characteristics of this kind form one basis for many if not all of the contrasts that are observed in languages. Humans are, in a sense, endowed with a set of contrasts that have evolved from these articulatory-perceptual relations. The phonological component of lexical items stored in the memory of a listener/speaker is presumed to be represented in terms of distinctive features that specify these contrasts.

These articulatory-acoustic-perceptual relations that are postulated to “define” the distinctive features are only observed if certain postural or default conditions for speech production are in place. For example, during speech production, the respiratory system operates in a particular way, with passive relaxation forces playing a large role in maintaining a subglottal pressure with only small variations [12]. The default configuration for the soft palate is a raised position, causing the passage to the nasal cavity to be closed. This soft palate posture allows a class of obstruent consonants to be introduced. The default condition of the vocal folds could be regarded as a position for phonation in vowels. Thus the articulatory gestures and their acoustic and perceptual consequences that provide evidence for the emergence of a set of distinctive features are only part of a theory or model of how speech is produced and perceived. These gestures must be superimposed on a set of postural conditions.

Nevertheless, in the production of an utterance consisting of a sequence of words or morphemes, one might suppose that the skeleton for the articulatory gestures used by a speaker is based on the articulations that are specified by the quantal articulatory-acoustic-perceptual relations. Thus in producing a word like teen, the articulatory gestures for /t/, with the features [+consonant; -continuant; +stiff vocal folds; +tongue blade], specify that a complete closure be made in the oral cavity, the tongue blade is used to make this closure, and there is stiffening of the vocal folds, as well as the default postural requirements just noted. Some adjustment of the stiffness of the vocal-tract walls may also be made during the consonant closure to facilitate the buildup of intraoral pressure and hence to inhibit glottal vibration. These gestures must be actualized in order to achieve the acoustic goals that are part of the definitions of the features. Realization of these acoustic goals requires that the gesture for a feature for a given segment be properly timed relative to gestures for other features of the segment, so that the gesture provides the appropriate acoustic pattern and so that the gesture for one feature does not mask the gesture for another.

It is well known, however, that the inventory of the gestural instructions is much more than the skeletal pattern provided by the feature-defining gestures in the planning stage for the utterance, together with the postural conditions already noted. In the inventory of gestures just described for the word-initial /t/, for example, there is no specification of how the tongue body should be positioned or should move, how the vocal folds should be configured, or where the tongue blade should be placed along the hard palate. Additional gestural instructions must be added to those specified by the defining articulatory patterns for the features. For example a fronting of the tongue body is implemented for alveolar stop consonants like /t/, and it has been argued that this tongue body position will strengthen or enhance the acoustic consequence of the basic feature-defining gesture for [+tongue blade], and will result in increased perceptual saliency of this place of articulation relative to other places [13]. And the voiceless consonant /t/ in pretonic position is produced with a spread glottal configuration that persists a few tens of milliseconds beyond the consonant release. This glottal spreading gesture can be viewed as enhancing the perception of the sound resulting from the vocal-fold stiffening gesture that defines the feature [+stiff vocal folds], i.e., it will strengthen the perception of a lack of glottal vibration. These enhancing gestures are not, however, feature-related, in the sense that they define a contrast. Thus, for example, the glottal spreading that leads to aspiration for /t/ does not define a contrast in English (although a gesture of this kind may be contrastive in other languages).

These and other examples show that the enhancing gestures that are introduced to increase the perceptual saliency of a feature in a particular phonetic and prosodic context can also introduce acoustic properties in addition to the ones that define the feature [14],[15]. These enhancing gestures and their acoustic consequences tend to be graded, and are not “quantal” in the sense of the feature-defining gestures and their acoustic and perceptual correlates. The enhancements that are used for particular features may differ from one language to another.

 ISBN 1-876346-48-5 © 2003 UAB

In the production of a word like hot, in which /t/ occurs in syllable-final position, there may be contexts in which the /t/ release cannot be observed, and consequently there is no burst to provide a cue that identifies the position of the tongue blade. However, the formant transitions in the immediately preceding vowel are evidence for tongue-body fronting, which, in the view proposed here, is an enhancing gesture. Thus the enhancing gesture provides an appropriate cue in the absence of the acoustic pattern that arises from the feature-defining gesture.

This and other examples show that the defining articulatory and acoustic patterns for a given feature are sometimes not directly observable in the acoustic signal due to gestural overlap. Nevertheless, the distinctive feature that defines a particular contrast may be represented in the sound by acoustic cues that are introduced to enhance the primary feature-defining cues.

In examining and interpreting the articulatory, acoustic, and perceptual aspects of the phonetic patterns that occur in a given language, it may not always be clear which are the properties that define the feature or the contrast and which are the enhancing properties. Both the defining and the enhancing properties are presumably used by a listener in uncovering the distinctive features that are assumed to be units in terms of which words are stored in memory. And in producing an utterance, the defining gestures presumably provide a skeleton for assembling the sequence of articulatory gestures, but these are embellished by enhancing gestures in a way that often depends on the context in which particular features occur.

In the study of phonetics, then, we have at least two kinds of tasks. There is the task of determining the articulatory, acoustic, and perceptual bases for the existence of the distinctive features that specify the contrasts that appear to exist in the languages of the world. And, in a given language, there is the problem of determining how these contrasts are instantiated by speakers --- what are the "postures" or default conditions for the articulatory and respiratory system, what are the enhancing gestures for given features in particular contexts, and what kinds of gestural overlap are tolerated.

5. ACKNOWLEDGEMENTS

This research was supported in part by grant DC00075 from the National Institutes of Health. Advice and technical help from Morris Halle, Jay Keyser, Helen Hanson and Janet Slifka is gratefully acknowledged.

REFERENCES

[1] R. Jakobson, C. G. M. Fant and M. Halle, "Preliminaries to speech analysis: The distinctive features and their correlates," Cambridge, MA: MIT Press, 1967.

[2] N. Chomsky and M. Halle, *The Sound Pattern of English*, New York: Harper and Row, 1968.

[3] B. Cranen and L. Boves, "On subglottal formant analysis," *J. Acoust. Soc. Am.*, vol. 81, pp. 734–746, 1987.

[4] D. H. Klatt and L. C. Klatt, "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *J. Acoust. Soc. Am.*, vol. 87, pp. 820-857, 1990.

[5] H. M. Hanson and K. N. Stevens, "Sub-glottal resonances in female speakers and their effect on vowel spectra," *Proc. XIIIth International Congress of Phonetic Sciences (ICPhS-95*), vol. 3, pp. 182-185, Stockholm, Sweden, 1995.

[6] K. N. Stevens, *Acoustic Phonetics*, Cambridge MA: MIT Press, 1998.

[7] M. Y. Chen, "Acoustic correlates of English and French nasalized vowels," *J. Acoust. Soc. Am.*, vol. 102, pp. 2360-2370, 1997.

[8] S. Maeda, "The role of sinus cavities in the production of nasal vowels," *Proc. ICASSP-82*, Paris, pp. 911-914, 1982.

[9] K. N. Stevens, "The quantal nature of speech: Evidence from articulatory-acoustic data," in *Human Communication: A Unified View*, P. B. Denes and E. E. David Jr., Eds., pp. 51-66. New York: McGraw Hill, 1972.

[10] K. N. Stevens, "On the quantal nature of speech," *J. Phonetics*, vol. 17, pp. 3-46, 1989.

[11] K. N. Stevens, "The properties of the vocal-tract walls help to shape several phonetic distinctions in language," *Travaux du Cercle Linguistique de Copenhague*, vol. XXI, pp. 285-297, 2001.

[12] J. Slifka, "Respiratory constraints on speech production at prosodic boundaries," Ph.D. thesis, Harvard-MIT Div. Of Health Sciences and Technology, MIT, 2000.

[13] S. J. Keyser and K. N. Stevens, "Enhancement revisited," in *Ken Hale: A Life in Language*, M. Kenstowicz, Ed., pp. 271-291. Cambridge MA: MIT Press, 2001.

[14] R. L. Diehl, "The role of phonetics within the study of language," *Phonetica* vol. 48, pp. 120-134, 1991.

[15] J. Kingston and R. Diehl, "Intermediate properties in the perception of distinctive feature values," in *Papers in Laboratory Phonology IV*, B. Connell and A. Arvaniti, Eds., pp. 7-27. Cambridge University Press, 1995.