Predicting the Structure of Phonological Systems

Patrice Speeter Beddor
Program in Linguistics, University of Michigan, Ann Arbor, Mich., USA

Abstract. Do the phonological predictions derived from phonetic models impose restrictions on the definition of 'possible phonological system'? This question is addressed by examining the spectral integration model of vowel perception. The output of the model is investigated under a series of phonetic manipulations and translated into a set of phonological predictions, which are then compared against cross-linguistic patterns in vowel processes and inventories. It is argued that, despite the serious constraints on predictive power imposed by language-specific aspects of phonetic and phonological structure, phonetic models such as the spectral integration model enable us to define the set of phonetically motivated phonological phenomena.

One of the goals of phonological theory is to delimit or define the set of possible phonological systems. Phonetic theory has, in recent years, focused on the development of models of the human speech mechanism. These two goals overlap in that the phonetic models have been used to generate sets of predictions concerning the structure of phonological systems. The present paper considers the possible consequences of this approach for phonological theory by asking whether the predictions derived from phonetic models impose restrictions on the definition of 'possible phonological system'.

In answering this question, this paper examines one specific phonetic model, the spectral integration model of vowel perception. The model is used to generate a set of predictions concerning the effects of five phonetic manipulations - nasal coupling, horizontal tongue body constriction, lip rounding, breathy voice, and creaky voice - on perceived vowel height. The general question of the consequences of phonetic models for phonological theory is then addressed in terms specific to this model: To what extent do the automatic constraints on vowel processing imposed by the spectral integration model translate into phonological predictions concerning the nature of vowel raising and lowering processes and the structure of vowel inventories?
1. Phonetic Modeling

Quantitative models of specific components of the speech process have many applications, one of which is in phonological systems. Interpreting a phonetic model in terms of its phonological consequences often involves some version of the following three steps: First, the response of the model is investigated under selected phonetic conditions, yielding a set of output conditions. Second, the output conditions are interpreted as a set of predictions concerning the structure or behavior of phonological systems. Third, these predictions are tested against such aspects of phonological systems as features, units, inventories, and processes. The focus in this paper is on predictions as tested against phonological inventories and processes. [Concerning the relation of phonetic models to phonological features and units, see, for example, Stevens, 1983 and 1989, and Browman and Goldstein, 1989, respectively, and also Keating, 1988, for an overview.]

Each of the three steps just described rests on a set of assumptions. The phonetic model itself is of course only an approximation to a component (e.g., the aerodynamic, articulatory, or acoustic component) of the speech process. Thus, in the first step, the validity of the output conditions of the model depends on the adequacy of both the model and the characterization of the input conditions. In the third step, it is the adequacy of the available language data that is potentially problematic. That is, a fundamental assumption is that language data in grammars or other linguistic accounts provide sufficiently detailed and accurate phonetic descriptions to serve as an appropriate test of the model’s predictions.

In the second step, phonetic models tend to be interpreted in one of two ways with respect to their phonological consequences. Under one interpretation, the output conditions of the model are viewed as physical limits which have fairly direct consequences for the behavior of speaker-hearers; these phonetic consequences may become incorporated into the phonological system. This approach is represented by the work of Ohala (e.g., 1974, 1981, 1983), who argues that physical constraints result in speaker-induced or listener-induced variation in the speech signal. Such variation may (under appropriate conditions not, for the most part, specified by phonetic models) result in sound change, which is synchronically reflected in phonological processes or the structure of phonological inventories. For example, calculations using models of speech aerodynamics indicate that voicing can be sustained longer in a front-articulated stop than in a back-articulated stop (Ohala, 1983; Keating, 1984). Phonetically, if preferred closure duration were the same for all places of articulation, back stops might be expected to sporadically co-occur. Systematic incorporation of such occasional co-occurrence into a language should result in allophonic or morphophonemic alternations or, if applied to all instances of a back stop, loss of that stop (e.g., /g/ in the segment inventory. Cross-linguistic phonological patterns support these predictions; devoicing processes appear to be favored in back stops and voiced stop inventories are more likely to lack back stops than anterior ones (Ohala, 1983; Maddieson, 1984).

Under a second interpretation of phonetic models, the output conditions are viewed as default or optimal settings (as opposed to physical limits). This section presents an adaptation to this point of view.
nd step, phonetic models tend to be in one of two ways with phonological consequences. Interpretation, the output condition model are viewed as physical have fairly direct consequences of speaker-hearing consequences may be abstracted into the phonological approach is described by the [e.g., 1974, 1981, 1983], who К physical constraints result in ed or listener-induced variation signal. Such variation appropriate conditions not, for specified by phonetic mod-sound change, which is syn- reflected in phonological pro-structure of phonological in-examples, calculations using each aerodynamics indicate can be sustained longer in a ed step than in a back-articula-tion step as required by the Keating, 1984], if preferred closure duration for all places of articulation, might be expected to sporad-ically Systematic incorporation of al devoicing into a language in allophonic or morphopho- tions or, if applied to all in-back step, loss of that step the segment inventory. Cross- onal patterns support ons: devoicing processes ap- in back stops and voiced es are more likely to lack back- erior ones [Ohala, 1983; Maddieson, 1988]

2. A Phonetic Model: Spectral Integration

This section presents the spectral integration model of vowel perception as it has been described and developed in the literature and considers its implications for phonological systems. The limitations of this model and approach need to be recognized. The spectral integration model is based on imperfect understanding of both acoustic-to-auditory transformations and the effect of these transformations on the perception of vowel quality. The model is also based for the most part on static characteristics of the vowel spectrum and does not take into account coarticulatory influences on perceived vowel quality. Further, by viewing the spectral integration model in isolation, other phonetic influences on vowel systems are being ignored. The approach taken here is to temporarily minimize these limitations in order to maximize the predictive strength of the model — in effect, to push the model as far as possible (sections 2.1 and 2.2). Phonetic and phonological restrictions on this approach are then discussed (section 2.3), leading to limited predictions which are tested against phonological data (section 2.4).

2.1. Description of the Model

The theory of spectral integration predicts that, when two adjacent peaks in a vowel spectrum are close in frequency, the perception of vowel quality is determined by a weighted average of this region of the spectrum rather than by the frequency of the individual peaks. This theory rests on a perceptual phenomenon first demonstrated by Delattre et al. [1952], who found that perceptually acceptable one-formant approximations to nonfront vowels could be synthesized by a single formant intermediate in frequency between $F_1$ and $F_2$ of the
Fig. 1. Schematic representation of the spectral integration phenomenon according to Chistovich's model. When two spectral peaks are within 3–3.5 Bark of each other, the frequency corresponding to the perceptual center of gravity (see text), represented by the dashed line, is an amplitude-weighted measure of the mean frequency of the two peaks.

corresponding naturally produced vowels. Front vowel approximations generally required two formants, with the preferred stimuli having a second formant higher in frequency than that of F2, of natural front vowels. Given the proximity of F1 and F2 in back vowels, and of F2 and F3 in front vowels, Delattre et al. [1952] speculated that listeners' responses were governed by formant spacing. This phenomenon of averaging or integration of spectral peaks close in frequency has been substantiated in numerous subsequent experiments e.g., Miller, 1953; Carlson et al., 1975; Bladon and Fant, 1978; Traunmüller, 1981; Beddor and Hawkins, 1990.

The detailed spectral characteristics under which spectral integration occurs have been investigated by Chistovich and Lublinskaya [1979] and Chistovich et al. [1979] and incorporated into a model of auditory spectrum processing [Chistovich, 1985]. The focus here is on those aspects of the model specifically related to spectral integration. Two aspects are particularly noteworthy. First, based on listener matches for vowel quality between one-formant and two-formant synthetic vowels, the frequency range within which spectral integration occurs was calculated at 3–3.5 critical bands of Bark. Adjacent spectral peaks are perceived as a single 'center of gravity' only when they fall within this critical range. Second, when two formants are within the critical distance, the perceptual center of gravity of that region of the vowel spectrum is influenced by relative formant amplitude. As schematically represented in Figure 1, the perceptual center of gravity (i.e., the frequency of the one-formant stimulus that listeners would choose as most similar to the two-formant reference vowel) falls in the middle of the spectral region when the two formants have equal amplitudes, but is closer to the greater-amplitude peak when the formants have different amplitudes. The model defines the centroid, an amplitude-weighted measure of mean frequency in the critical frequency range, as a spectral measure of the effective perceptual center of gravity, according to the formula [based on Chistovich, 1985]

\[ X_{cen} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} Y_i} \]

where X = frequency (Bark) and Y = response magnitude (variably computed in linear, power, and log scales).

Thus Chistovich's model predicts that spectral smoothing applies within a band of 3–3.5 Bark. The critical distance of roughly 3 Bark is well attested [Bladon, 1983; Sydall, 1985]. However, recent studies suggest the need to modify Chistovich's formulation of the smoothening function in terms of the relative weighting of formant frequency and amplitude. Klatt [1985] covaried formant amplitude and frequency such that the perceptual center of gravity should remain constant according to formulation and found a reduction of formant frequency: vowel quality outweighs relative amplitude. Beddor and Hupke [1988] noted a critical band effect for formant bands that formant frequency shifted vowel quality when widths were narrow, while vowel shape exerted more influence on wider bandwidths.

Based on these findings, a model adopted here assumes the function of the 3–3.5 critical frequency-weighted frequency relative weight assigned increases as spectral prominence. While a precise spectral integration function is to be determined, the data suggests a relation of a shift in the perception of formant frequency that results from manipulation, and in this case can be investigated under different conditions.

2.2. Output of the Model

The low-frequency component

This section investigates the spectral correlates of the perceived vowel height dimension of height is correlated with the frequency of perceived vowel height increases e.g., Ney, 1952; Fant, 1960; 1 spectral integration modulates the output of the velar a preceding description, which works whose low-frequency sp...
ic vowels, the frequency range of spectral integration occurred at 3-3.5 critical bands or at spectral peaks are perceived 'center of gravity' only when near this critical range. Second, formants are within the critical perceptual center of gravity of the vowel spectrum is influential formant amplitude. As represented in figure 1, the center of gravity (i.e., the frequency of the one-formant stimulus that is closest as most similar to the reference vowel) falls in the spectral region when the two equal amplitudes, but a greater-amplitude peak when different amplitudes are used. The centroid, an amplitude measure of mean frequency in the frequency range, as a spectral more-effective perceptual center according to the formula [based on Chistovich 1985]

$$\frac{\sum Y_i}{\Sigma Y}$$

frequency (Bark) and $Y_i$ = re-

nitude (variably computed in $r$, and log scales). Chistovich's model predicts that nothing applies within a band of the critical distance of roughly 125 Hz in this range. However, recent studies suggest that Chistovich's formula smoothing function in terms of weighting of formant frequency. Klatt [1985] covered for-}

nitude and frequency such that all center of gravity should remain constant according to Chistovich's formulation and found that the contribution of formant frequency to perceived vowel quality outweighed that of formant amplitude. Beddor and Hawkins [1990] manipulated formant bandwidth and found that formant frequency dominated perceived vowel quality when formant bandwidths were narrow, whereas overall spectral shape exerted more influence when bandwidths were wide.

Based on these findings, the model adopted here assumes that the integration function of 3-3.5 critical bands is an amplitude-weighted frequency measure in which the relative weight assigned to amplitude increases as spectral prominence decreases. While a precise spectral measure of the integration function is lacking, the description is sufficiently detailed to predict the direction of a shift in the perceived center of gravity that results from a given spectral manipulation, and in this respect the model can be investigated under a variety of phonetic conditions.

2.2. Output of the Model: Manipulating the Low-Frequency Center of Gravity

This section investigates the consequences of the spectral integration model for perceived vowel height. The perceptual dimension of height is often described as correlating with the frequency of $F_1$, such that perceived vowel height lowers as $F_1$ frequency increases [e.g., Peterson and Barney, 1952; Fant, 1960; Lindau, 1978]. The spectral integration model as presented above leads to the expectation that, while this description works well for vowels whose low-frequency spectrum is dominated by $F_1$, a more general characterization is that perceived height correlates with the low-frequency region of spectral prominence. Under this view, phonetic manipulations which shift the center of gravity of the low-frequency region of prominence in a vowel spectrum are predicted to influence perceived vowel height. Below, the specific consequences of five such manipulations are considered: nasal coupling, horizontal tongue body position and lip position (taken together as oral cavity shape), and breathy and creaky voice (taken together as phonation type).

2.2.1. Nasal Coupling

Coupling the nasal tract to the oral tract in the production of a nasal vowel principally affects the vowel spectrum in the low frequencies, where a pole-zero pair is added in the vicinity of $F_1$ [Fujimura and Lindqvist, 1971; Stevens et al., 1987]. A consequence of the nasal pole-zero pair is that the low-frequency spectral energy has a low amplitude and is distributed over a broad frequency region relative to the spectrum of the corresponding nonnasal vowel [House and Stevens, 1956; Hawkins and Stevens, 1985]. These changes in spectral shape are generally accompanied by frequency shifts in the center of gravity of this region of the spectrum. The frequency of the first 'oral' formant of the nasal vowel, $F_1$, is shifted relative to $F_1$ frequency of the corresponding nonnasal vowel. The frequency of the added nasal formant, $F_1$, is generally less than that of $F_1$ in low vowels, but greater than $F_1$ frequency in high and (higher-)mid vowels [Fant, 1960; Fujimura and Lindqvist, 1971]. These consequences of nasal coupling are illustrated in figure 2 by the vocal tract transfer functions of oral and
nasal versions of \( [a] \) and \( [i] \) generated on the Haskins Laboratories' articulatory synthesizer [described by Rubin et al., 1981]. In both nasal vowels, \( F_N \) is within 3.5 Bark\(^1\) of \( F_i \), but in the low vowel, addition of a moderate degree of coupling shifted the center of gravity of the low-frequency region slightly downward, while in the high vowel, coupling shifted the center of gravity upward. So, in general, \( F_N \) is predicted to lower the low-frequency center of gravity of low vowels and to raise it in high and mid vowels.

That these shifts in the center of gravity due to nasal coupling correlate with shifts in perceived vowel height has been demonstrated experimentally. Using formant syn-

---

\(^1\) Here and elsewhere in the paper, Bark values were calculated from Hertz values by the equation provided by Schroeder et al. [1979].
thesis, Beddor and Hawkins [1990] generated vowel sets in which a nasal vowel standard was paired with a series of oral vowels. The stimuli in each oral vowel series varied in the frequency of $F_1$, the perceptual effect being variation in vowel height. Listeners, who selected from each set the oral vowel which sounded most similar to the nasal standard, tended not to choose the oral vowel in which $F_1$ frequency was the same as $F_1$ frequency in the nasal vowel (except for [I], in which $F_1'$-$F_N$ distance exceeded 3.5 Bark). As predicted by the model, listeners preferred oral vowels in which $F_1$ frequency fell between $F_N$ and $F_1'$ of the nasal vowel standard.

While the general prediction of the model is that nasализation will lower the perceived height of high and mid vowels and raise that of low vowels, the model can also be shown to predict that the height consequences will differ for back versus front vowels. Because of the proximity of $F_1$ and $F_2$ in back but not front vowels (i.e., the low-frequency region of prominence in back nasal vowels often includes not only $F_1$ and $F_N$ but also $F_2$), it is not unlikely that the magnitude of the center of gravity shift due to nasализation will be less in back vowels, and the shift in perceived height will be correspondingly smaller. Given that current spectral measures do not accurately describe the perceptual center of gravity of the low-frequency region, this suggested difference is difficult to substantiate. But taking the centroid calculation as a rough description, centroid measures of the spectra of naturally spoken vowels do indicate this front-back asymmetry, the centroids of non-low front — but not back — nasal vowels being consistently higher than that of corresponding oral vowels [Beddor, 1982].

2.2.2. Oral Cavity Shape: Rounding and Backness

Although lip rounding and front-back tongue body constriction are separable phonetic manipulations, they are considered together here as they interact in their effect on formant proximity in the low frequencies of the vowel spectrum. The principal spectral consequence of lip rounding is to lower formant frequencies, especially $F_2$—$F_4$ [Fant, 1960; Ladefoged, 1982]. Horizontal tongue body constriction correlates with $F_2$ frequency, such that a back position leads to a maximally low $F_2$ and a front position to a maximally high $F_2$ [Fant, 1960; Stevens, 1972, 1983]. When lip rounding and back tongue position co-occur, as they do in most vowel systems of the world's languages, $F_1$ and $F_2$ are in relatively close proximity, with $F_2$ having a wide bandwidth [Stevens, 1989].

The clustering of $F_1$ and $F_2$ in back rounded vowels enhances perceived vowel backness [e.g., Linder, 1978; Ladefoged, 1982; Stevens et al., 1986]. $F_1$—$F_2$ proximity is also predicted, according to the spectral integration model, to influence perceived vowel height: for a vowel with a given $F_1$ frequency, perceived height should be lower when the $F_2$—$F_1$ distance is within 3.5 Bark than when it is not. This theoretical prediction is consistent with the empirical findings of Chistovich and her colleagues, although the perceptual consequences of their manipulations of $F_2$—$F_1$ proximity are described generally in terms of overall vowel quality. It is more directly supported by Beddor and Hawkins [1990], who describe the acoustic manipulations as influencing perceived height.

If increased proximity of $F_2$ to $F_1$ lowers perceived vowel height, then phonetic ma-
Manipulations which affect this proximity, such as lip rounding or unrounding and tongue body backing of fronting, might be expected to influence not only perceived roundness or backness, but also to shift perceived height. This claim has not been investigated experimentally, but acoustic measures of naturally spoken vowels show that linguistic use of rounding and backness have the expected effect on formant proximity. For example, acoustic measures of high back rounded [u] and unrounded [u] in Turkish show that $F_1$ and $F_2$ are generally within about 3 Bark for [u], but are separated by about 7 Bark for [u] [based on Selen, 1979; and Ergenç, 1989]. Manipulation of lip rounding by means of articulatory synthesis has a similar, albeit smaller, effect on $F_2-F_1$ separation. The vocal tract configurations corresponding to the transfer functions in Figure 3 differ only in lip protrusion, the main spectral difference being in $F_2$ frequency. Lip protrusion reduced $F_2-F_1$ separation to within the critical distance, which raised the low-frequency center of gravity, and is predicted to lower perceived vowel height.

Similarly, languages which contrastively manipulate the front-back dimension in vowels appear to exhibit the critical separation or proximity of formant frequencies. For example, some speakers of French maintain a distinction between two low unrounded vowels, central [a] and back [o]. Measures of formant frequencies show that $F_1$ and $F_2$ are within 2.1–2.9 Bark of each other for [a], but are 3.4–5.6 Bark apart for [o] [Debrock, 1974; Mettas, 1979]. Although these French vowels may also differ somewhat in height and lip rounding, these factors were controlled for in articulatorily synthesized versions of [a] and[o] represented by the transfer functions in Figure 4. Fronting of the tongue body center from [a] to [o] caused the $F_2-F_1$ separation to exceed the critical distance, so that in [o] the low-frequency center of gravity is determined exclusively by $F_1$ (as opposed to a weighted $F_1-F_2$ average in [a]), which is predicted to raise perceived vowel height.

In contrast, the addition of lip rounding to front vowels, or the backing of front vowels to a more central configuration, does not generally place $F_1$ and $F_2$ within the critical distance for spectral $F_2-F_1$ distance for the front of Swedish [Fant, 1969; French [Mettas, 1979], for 6.0 Bark. This distance is in central vowels, but is still in, for example, rounded vowel [e] of [Håkansson, 1982]. Therefore, based on that perceived height will low-frequency center of spectral integration may variation in lip rounding tongue body position perceived height of back vowel front vowels. The situa action, is more complex, predicted to have little producing vowels but not low vowels (due to reality) and rounded relatively low $F_2$ frequency.

2.2.3 Phonation Type: Creaky Voice
The phenomenon of creaky voice occurs not only when th...
Fig. 3. Transfer functions of articulatorily synthesized high back vowels differing in lip protrusion. F₂–F₁ distance is 5.9 Bark in the spread vowel (solid line) and 3.3 Bark in the rounded vowel (dashed line).

Fig. 4. Transfer functions of articulatorily synthesized low unrounded vowels differing in tongue body backness. F₂–F₁ distance is 3.0 Bark in the back vowel (solid line) and 4.2 Bark in the central vowel (dashed line).

languages which contrastively the front-back dimension in to exhibit the critical separa arium of formant frequencies, some speakers of French distinction between two low aels, central [a] and back [a]. Formant frequencies show that within 2.1–2.9 Bark of each but are 3.4–5.6 Bark apart for 1974; Mettas, 1979]. Although vowels may also differ somewhat in lip rounding, these are controlled for in articulatorily versions of [a] and [a] ree transfer functions in figure 4. The F₂–F₁ separation to exceed stance, so that in [a] the latter of gravity is determined F₁ (as opposed to a weighted t in [æ]), which is predicted to d vowel height.

2.2.3 Phonation Type: Breathy and Creaky Voice
The phenomenon of spectral integration occurs not only when two formants are in close proximity, but more generally when adjacent spectral peaks are within the critical range. Of particular importance is that spectral averaging has been demonstrated for the first harmonics, F₀, and F₁ (Fujisaki and Kawashima, 1968; Carlson et al., 1975), and F₂–F₁ proximity has been shown to influence perceived vowel height (Traunmüller, 1981).

The phonatory manipulations considered here are breathy voice, in which the vocal folds are abducted relative to unmarked or modal voicing, and stiff or creaky voice, in which the vocal folds are relatively abducted. Based on a quantitative model of the vocal folds, Stevens (1988) reported that the low-frequency spectral consequences of breathy phonation are increased bandwidth of F₁ and increased amplitude of F₁, resulting in F₀ amplitude being substantially greater than F₁ amplitude. In contrast, the vocal fold configuration in creaky voice (Stevens’ ‘pressed’ voice) leads to a decreased bandwidth of F₁ and increased amplitude of F₁ relative to that of F₀.

These predicted spectral effects have been found in naturally produced vowels of
languages in which breathiness or creak has a contrastive function. For example, in Jalapa Mazatec, which contrasts modal, creaky, and breathy vowels, F₀ amplitude is less than F₁ amplitude in creaky vowels and more than F₁ amplitude in breathy vowels, with the general finding being that the value (in dB) of the amplitude difference between F₀ and F₁ increases from creaky to modal to breathy voice [Ladefoged et al., 1988]. This generalization also holds for the breathy–nonbreathy distinction in Gujarati [Fischer-Jorgensen, 1967], Xuoo [Ladefoged and Antoñanzas-Barroso, 1985] and Dirka [Denning, 1989]. The Dirka contrast as produced by a speaker of the Bor dialect.

is illustrated by the FFT spectra of nonbreathy [e], F₀ amplitude (the third highest amplitude) is greater than F₁ amplitude (the second highest amplitude), and F₂ amplitude (the highest amplitude) is less than both F₀ and F₁. The third highest amplitude is less than both F₀ and F₁. Within the spectral int the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.

The predicted shifts in the effect of these many factors on the perceived height of breath. The predicted to be higher for breathy counterparts and the vowels to be lower for creaky counterparts.
is illustrated by the FFT spectra in figure 5. In nonbreathy [ɛ], $F_0$ amplitude is less than $F_1$ amplitude (the third harmonic carrying most of the energy due to $F_1$); this amplitude relation is reversed in breathy [ɛ].

Within the spectral integration model, the effect of these manipulations is for breathy voice to lower, and for creaky voice to raise, the center of gravity of the low-frequency region. (Possible accompanying differences, such as changes in pitch are being ignored. While such differences sometimes co-occur with phonation differences in languages, they need not.) Consequently, the perceived height of breathy voice vowels is predicted to be higher than their non-breathy counterparts and that of creaky voice vowels to be lower than their non-creaky counterparts.

The predicted shifts in perceived height depend on the assumption that $F_0$ and $F_1$ are within an integration band of 3–3.5 Bark. However, $F_0$–$F_1$ spacing does not satisfy this condition for all vowel heights and pitch ranges. (For example, given a vowel with an $F_0$ of 130 Hz, $F_1$ is within 3.5 Bark only if it has a frequency of less than 500 Hz.) This suggests that the predicted effect of breathiness and creak on perceived height is, in general, restricted to nonlow vowels. On the other hand, even in low vowels, phonation type influences $F_1$ bandwidth (which is broad in breathy vowels and narrow in creaky vowels) and $F_1$ bandwidth can affect perceived vowel height [Beddor and Hawkins, 1990]. The effect of increased bandwidth appears to be to increase the influence of the spectral envelope (as opposed to peak frequency) on perceived height. The prediction of the spectral integration model concerning height shifts in low nonmodal vowels therefore depends in part on the shape of the spectrum in the $F_1$ region. Whether breathiness or creak has a consistent effect with respect to this spectral property is unclear from the literature and hence, based on current descriptions, there are no predictions concerning the direction of shifts in the perceived height of low vowels due to phonation type.

### Table 1. Predicted height effects of changes in oral cavity shape

<table>
<thead>
<tr>
<th>Tongue body backness</th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retraction</td>
<td>-</td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td>Fronting</td>
<td>-</td>
<td>(round or low only)</td>
<td></td>
</tr>
<tr>
<td>Rounding</td>
<td>-</td>
<td></td>
<td>↑</td>
</tr>
<tr>
<td>Unrounding</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For each manipulation, ↑ indicates a predicted raising effect, ↓ indicates lowering, and – indicates that no height shift is predicted.

### Table 2. Predicted height effects of nasalization and phonation type

<table>
<thead>
<tr>
<th>Tongue body height</th>
<th>high / mid</th>
<th>low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasalization</td>
<td>↓</td>
<td>↑</td>
</tr>
<tr>
<td>Breathiness</td>
<td>↑</td>
<td>?</td>
</tr>
<tr>
<td>Creak</td>
<td>↑</td>
<td>?</td>
</tr>
</tbody>
</table>

Symbols are as in table 1; ? indicates that the direction of the expected height shift is unclear from available acoustic data.

2.2.4 Summary of the Height Shifts

Predicted by the Model

The spectral integration model has been interpreted here as predicting that a shift in
the low-frequency center of gravity will influence perceived vowel height. Based on what is known about the consequences of nasal coupling, oral cavity shape, and phonation type in the low frequencies of the vowel spectrum, it is predicted that these manipulations will shift perceived height as summarized in tables 1 and 2. The height effects of oral cavity shape (table 1) depend on vowel backness, with effects being restricted to back vowels (and to a subset of central vowels in the case of retraction). On the other hand, the height effects of nasalization and phonation type (table 2) are primarily conditioned by vowel height, with the height effects of nasalization being more pronounced in front vowels.

2.3. Translation into Phonological Predictions

It is emphasized that, if we accept the model's representation of vowel processing, the predicted shifts in perceived height automatically follow. (The exact mechanisms of this automaticity are not directly relevant to this paper.) How, then, are these automatic consequences of vowel processing translated into predictions for phonological systems?

2.3.1. Articulatory Consequences of Perceptual Constraints

The translation here of perceptual constraints into phonological predictions follows a view of the contribution of the listener to phonological change (and hence synchronic phonological patterns) that is implicit in the work of Sweet [1888] and Paul [1891], and made explicit by Ohala [e.g., 1981, 1989]. Within the context of height effects, the view is that, if a listener (presumably, the language learner) perceives a vowel as being relatively high or low—for example, if a breathy voice vowel is heard as higher than its nonbreathy equivalent or a nasal vowel is heard as lower than its oral equivalent—then the learner will in turn produce that vowel with a relatively high or low tongue body configuration. That is, the shift in perceived vowel height that results from a shift in the low-frequency center of gravity is reinterpreted by the learner in terms of tongue body height [see Ohala, 1981, and Beddor et al., 1986, for further discussion]. Thus the interpretation here of the spectral integration model with respect to its phonological consequences parallels the first approach described in section 1.2 That is, the constraints on vowel processing imposed by the model are interpreted as perceptual limits having direct consequences for the behavior of speaker-listeners—consequences which become incorporated into the phonological system.

The question of how the shifted vowel articulations of individual speaker-listeners are incorporated into the phonological system as a stable pattern is in large part a sociolinguistic one. However, if the perceptual height effects of such processes as, say, breathy voice and nasalization are indeed automatic consequences of vowel processing...

2 This is not to say that the spectral integration phenomenon is not compatible with the second interpretation. For example, the spectral integration model could be viewed as generating auditory vowel spectra that serve as input to a perceptual distance metric [e.g., Lindblom, 1986]. However, this approach involves additional considerations that go beyond the scope of this paper.
The view is that, if a listener-language learner perceives breathy or nasal vowels as oral equivalents—then the term produce that vowel with or without vowel, the shift in perceived vowels from a shift in the low-frequency center of gravity is reinterpreted in terms of tongue body and, 1981, and Beddor et al., 1983, discussion. Thus the integration of the spectral integration vector to its phonological cells the first approach done. That is, the constraints are perceived limits having consequences for the behavior of certain morphological correspondences which exist into the phonological of how the shifted vowel individual speaker-listeners into the phonological system is in large part as said. However, if the perception of such processes as, say, nasalization are indeed consequences of vowel processing, then we would expect shifted articulations of breathy or nasal vowels (at least in the speech of language learners) to be the norm rather than sporadic variants. Under this view, it would appear that the output of the spectral integration model should translate into strong predictions concerning vowel raising or lowering patterns in phonological systems, such that languages with breathy vowels, nasal vowels, and so on, are expected to exhibit the height patterns described above.

2.3.2. Limitations on the Consequences of Auditory Constraints for Phonological Systems

But such expectations are not satisfied by cross-linguistic study of the relevant phonological patterns. For example, Ruhlen's (1978) survey of 155 languages with distinctive nasal vowels found no reported differences (other than nasality) between the oral and nasal vowel inventories of half of the languages. While the inventories of many of the remaining languages did exhibit height differences (and while some of the other inventories may have had height differences that went unreported), it seems that height shifts in phonological systems are not necessary consequences of vowel nasalization. However, given that nasalization can be shown to influence perceived height under a variety of experimental settings (Wright, 1986; Krakow et al., 1988; Beddor and Hawkins, 1990), there would appear to be a discrepancy between the automatic constraints imposed by the spectral integration model and the phonological instantiation of these constraints. Put another way, the strength of our phonological predictions appears to be weaker, in terms of the likelihood of occurrence in a given phonological system, than might be expected on the basis of laboratory findings.

Many factors may account for this difference. One general factor which potentially influences whether the constraints imposed by the model are manifested in a given phonological system is phonetic variation. Languages may differ, for example, in the size of the shifts in the low-frequency center of gravity caused by changes in nasalization, oral cavity shape, and phonation type. Consequently, in some languages, the center of gravity shift might not be of sufficient magnitude to effect a phonological height shift. That is, if the perceived differences are quite small, shifted articulations on the part of language learners are expected to be more sporadic, and the likelihood of the shift becoming a stable phonological pattern is lessened. For example, the combined effects of backness and rounding result in an F2-F3 separation of less than 3.5 Bark in Turkish /u/, but this separation exceeds the critical distance in American English /u/ [Syrdal, 1985]. (This difference is consistent with the informal observation that Turkish /u/ is more rounded than that of American English.)

Languages apparently differ not only in the magnitude but also in the nature of the acoustic consequences of manipulations that are described as rounding, breathiness, and so on. For example, the vowel systems of the related West Nilotic languages of Dinka and Acholi are both described as having a breathy–nonbreathy voice distinction, but Denning's [1989] acoustic analyses of 2 speakers of each of these languages showed that the expected F2-F3 amplitude differences were present only in the Dinka vowels. Physiological studies of the two languages might explain this result, but of im-
importance here is that what is labeled breathy voice in these languages has different acoustic manifestations, leading to different predictions as derived from the model.

Phonetic variation introduced by individual speakers of a given language can also have critical consequences for the model's predictions. Figure 6 shows FFT spectra of two tokens of [e] produced by 2 speakers of Hindi. $F_1-F_N$ spacing is within the critical distance only for the vowel of speaker A. Analysis of ad [e] showed this to be a co between these speakers; was within 3.5 Bark for s and exceeded 4.5 Bark once again, these different predictions: [e] should lower than [e] when produced but not speaker B.

A second general fact the potential phonologic cephalic constraints is that to compensate for certainty cephalic consequences is linked to the phonological native language. The data manipulation which is the center of gravity vowel height assumes instances, listeners may also manipulate the height of the vowel in such a way as to produce a more spectral effect. For the most part, isolated voice is a phonetic concept judged relative to their context [Lindblom and 1967; Strange et al., 1987].

The nature of the cue suggests that listeners' spectral patterns enables assess the articulatory source of a given environment. The difference in $F_1-F_N$ spacing is at least in part to the voiced nasalized [e] as determined perceptual speaker B.
Analysis of additional tokens of [e] showed this to be a consistent difference between these speakers: $F_1 - F_N$ distance was within 3.5 Bark for speaker A’s vowels and exceeded 4.5 Bark for speaker B’s.\(^3\) Once again, these differences lead to different predictions: [e] should be perceptually lower than [e] when produced by speaker A, but not speaker B.

A second general factor that influences the potential phonological effects of perceptual constraints is that listeners appear to compensate for certain predicted perceptual consequences in ways possibly linked to the phonological structure of their native language. The claim that a phonetic manipulation which shifts the low-frequency center of gravity affects perceived vowel height assumes that, in some instances, listeners may exhibit imperfect knowledge of the spectral consequences of that phonetic manipulation. This claim is based on the most part on listener responses to isolated vowels. Vowels embedded in a phonetic context, however, are judged relative to their coarticulatory context [Lindblom and Studdert-Kennedy, 1967; Strange et al., 1976; Fowler, 1981], and the nature of the contextual influence suggests that listeners’ sensitivity to coarticulatory patterns enables them to correctly assess the articulatory configuration which generated the acoustic pattern. How this sensitivity comes about is controversial, but when a listener succeeds in identifying the articulatory source of a given vowel spectrum, there will be no shift in articulation when the listener produces that vowel, and hence no potential for phonological change [Ohala, 1981, 1986].

Listeners apparently compensate for spectral integration in certain circumstances. For example, the perceptual data reported in Krakow et al. [1988] indicate that, under appropriate conditions, listeners adjust for the effects of a shift in the low-frequency center of gravity on perceived vowel height. In that study, listeners identified articulatorily synthesized oral vowels (in a [bVd] context), noncontextually nasalized vowels ([bVd]), and contextually nasalized vowels ([bVnd]). The low-frequency spectral effect of velar port opening was to raise the center of gravity in the nasal vowels relative to that of the corresponding oral vowel. American English subjects’ judgments showed that, when comparing vowels generated with the same tongue configuration, the perceived height of oral and contextually nasalized vowels was the same, but that of noncontextually nasalized vowels was lower, indicating that the perceived height of nasal vowels is not entirely a function of their low-frequency center of gravity. When listeners are provided with an appropriate context for vowel nasalization, as in the [bVnd] context for English, they are apparently able to correctly assess the relative contribution of nasal coupling and tongue body configuration to perceived vowel height. It is expected that similar restrictions apply to the perceived height effects of the other phonetic manipulations discussed above, and hence whether these manipulations result in phonological shifts in vowel height in a given language should depend in part on the phonological structure of that language.
2.3.3. The Nature of Phonological Predictions Based on the Spectral Integration Model

It follows from the limitations just considered that, although the phenomenon of spectral integration imposes automatic constraints on vowel processing, the consequences of this phenomenon for perceived vowel height do not necessarily entail consequences for a given phonological system. The first factor (i.e., phonetic variation in parameters influencing the center of gravity) shows that strong phonological predictions would require detailed descriptions of spectral structure that are generally not available. Although such descriptions could, in principle, be obtained through appropriate phonetic analyses, the second factor (i.e., listeners' perceptual strategies) shows that even detailed spectral descriptions would not be sufficient, as spectrally identical vowels can elicit different percepts depending on the context in which they occur. Such problems in the translation of the model's output into phonological consequences are consistent with certain of the arguments offered by Ladefoged [1984, 1989] concerning the limitations of phonetic explanation.

Given the current understanding of spectral integration and the factors contributing to perception of vowel height, and the phonetic and phonological limitations just discussed, the spectral integration model cannot be interpreted as predicting that all languages will manifest height effects of spectral integration, nor which languages will manifest such effects. However, as interpreted here, the spectral integration model both predicts that languages will exhibit correlations between nasalization, tongue body backness, rounding, breathiness, or creak on the one hand and tongue body height on the other, and specifies the nature of these correlations. That is, the model is interpreted as predicting, for languages exhibiting a shift in vowel height correlating with one or more of these properties, the direction of that shift. The final step in the application of the spectral integration model to phonological systems is to test these predictions against phonological data.

2.4. Phonological Verification

Ideally, phonological testing of the model's predictions should involve selection of a sufficiently large, genetically balanced language sample and analysis of the vowel inventories and processes (and, if available, historical vowel shifts) of the selected languages in terms of the height effects of nasalization, oral cavity shape, and phonation type. But phonological descriptions differ considerably in purpose and theoretical orientation, and only a subset provide detail sufficient for these purposes. In practice, then, the size and balance of such surveys are severely limited by the scarcity of appropriate sources [for further discussion of this problem see, e.g., Maddison, 1984]. The survey presented here is even more limited in that investigation of the phonological consequences of the spectral integration phenomenon is an ongoing project and hence preliminary in some areas.

For the reasons discussed in the preceding section, phonological testing of the spectral integration model is restricted to languages in which vowel height appears to interact with nasalization, backness, rounding, breathiness, or creak. Such testing yields both supporting and conflicting data. In the interests of brevity, language-specific data are presented; rather, the results in terms of general consequences of vowel raising and nasalization are presented.

2.4.1. Height Effects of Nasalization

The phonological effect of nasalization on vowel height is related to several variables. Beddor [1982] analyzed the phonology of morphophonemic processes of vowel raising and lowering and found that oral-nasal nasal positioning has the following effects: (a) Nasalization raises the height of vowels, and (b) Nasalization lowers the height of front vowels. The general conclusion is that nasalization raises the height of front vowels. This conclusion is consistent with the nasalization shifts reported by Bhat [1975] and the vowel inventory differences between languages and dialects.

The phonological data on nasalization is based on a study in two respects: the height dimension of vowel nasalization (i.e., lowering of height and raising of low vowels) and the nasalization for mid vowels is not predicted by the model. The raising effect of nasalization may be linked to temporal factors.
...one hand and tongue body to other, and specifies the nature of relations. That is, the model is predicting, for languages exhibiting a shift in vowel height, correlating more of these properties, the drift shift. The final step in the so called spectral integration model is to test these predictions by phonological data.

**Logical Verification**

Phonological testing of the predictions should involve sufficiently large, genetically balanced samples and analysis of the sources and processes of historical vowel shifts of the form in terms of the height of oral cavity, backness and rounding. But phonological descriptions, in purpose and orientation, are only a subset of sufficient for these purposes, even when the size and balance of the observations is sufficiently limited by the sources of evidence.

The survey presented here is the first in that investigation of the consequences of the speech phenomenon is an ongoing research topic.

### 2.4.1. Height Effects of Nasalization

The phonological effects of nasalization on vowel height are relatively well-established due to several cross-linguistic surveys. Beddor (1982) analyzed allophonic and morphophonemic processes of nasal vowel raising and lowering in 75 languages and found that oral-nasal height differences reflected the following tendencies: (a) Nasalization raises the height of low vowels, (b) Nasalization lowers the height of high especially front vowels, (c) Distinctive nasalization lowers the height of mid especially front vowels, (d) Allophonic nasalization raises the height of mid especially back vowels. These patterns are generally consistent with the diachronic height shifts reported by Schourup (1973) and Bhat (1975), and the oral and nasal vowel inventory differences discussed in Ruhlen (1978).

The phonological data closely parallel the predictions based on the spectral integration model in two respects: the contraction of the height dimension due to nasalization (i.e., lowering of high and mid vowels and raising of low vowels) and the more pronounced lowering effect of nasalization on front than on back vowels. A third aspect, the different consequences of distinctive as opposed to allophonic or contextual nasalization for mid vowel height was not predicted by the model.

The raising effect of contextual nasalization may be linked to temporal factors. Historically, phonemicization of vowel nasalization is usually accompanied by loss of an adjacent nasal consonant and compensatory vowel lengthening [de C hemen and Anderson, 1979; Clements, 1982]. Contextual nasal vowels may therefore be shorter than their noncontextual counterparts and experimental evidence indicates that short vowels are perceived as higher [Krakow et al., 1988] and less nasal [Whalen and Beddor, 1989] than the corresponding long vowels. Alternatively, Ohala (1986) has suggested that the raising effects of contextual nasalization might be due to listener overcompensation. Although the presence of a nasal consonant normally should enable a listener to correctly assess the contribution of nasal coupling to perceived vowel height, the listener might overestimate the contribution of nasal coupling, thereby raising vowel height [see Krakow et al., 1988, for evidence consistent with this view as well].

### 2.4.2. Height Effects of Oral Cavity Shape

In a preliminary effort to determine whether phonological patterns of vowel raising and lowering are linked to vowel rounding or backing, I inspected the vowel inventories of the 317 languages in the UCLA Segment Inventory Database (UPSID) as reported by Maddison (1984). If the lowering effect of rounding on the perceived height of back vowels predicted by the spectral integration model has phonological consequences, then languages with a rounding contrast in back vowels should exhibit asymmetries such that the rounded vowels are lower than their unrounded counterparts. Similarly, if tongue body retraction lowers the perceived height of nonfront (rounded or low) vowels, then
vowel inventories with a central-back contrast should also show height differences, with the central vowels being higher than their back equivalents.

The first hypothesis of a correlation between rounding and height was not supported by the UPSID vowel inventories. Of the 27 languages which had both back rounded and unrounded (nonlow) vowels, the majority had rounded and unrounded vowels of the same height. In the 11 languages which showed height asymmetries in the back vowel system, there was no obvious pattern in the height differences attributable to lip rounding.

The hypothesis that backness correlates with height fared somewhat better. Only 9 languages in the UPSID sample had plain central rounded vowels, and 5 of these had back rounded vowels of the same height as the central ones. But in the other 4 languages, the back vowel was lower than its central counterpart (e.g., both Mongolian and Somali have high central /u/ as opposed to lower-high back /o/).

Assessment of the phonological height effects of rounding or backness is limited using this approach given the relatively small number of languages having back unrounded or central rounded vowels. Analysis of the combined effects of rounding and backness offers more compelling results. Since both lip rounding and retracted tongue position are predicted to lower vowel height, comparison of back rounded and front unrounded vowel systems should show the expected height differences. Such comparison encompasses over 98% of the UPSID languages. In the vast majority of these languages, the height of the highest back rounded vowel was the same as that of the highest front unrounded vowel (i.e., /u/ and /i/). However, when the height differed, a lower back vowel was nearly 4 times more likely than a lower front vowel. Specifically, the highest front unrounded vowel was lower than the highest back rounded vowel in 8 languages, the pattern in all 8 being /u/ lower. Twenty-nine languages had the reverse height difference, with the patterns including relatively low back vowels (i.e., not only /a/ but also /i/ and /o/). The frequent absence of /u/ relative to /i/ in vowel systems has been previously noted (e.g., Crothers, 1978; DiSci, 1984).

The phonological changes described in Labov et al. [1972] also suggest an influence of vowel backness and rounding on vowel height. Their analysis of ongoing change shifts in the vowels of various English dialects, combined with study of completed shifts in other languages, showed that the fronting of back tense (rounded) vowels in such shifts may be accompanied by a raising movement. In contrast, front vowel chain shifts, which primarily involve the height dimension, do not show a height-backness (or height-rounding) interaction. Other phonetic factors, of course, may be involved here; see Goldstein, 1983, for another view.

Thus preliminary data from vowel inventories and historical change are generally consistent with the claim that the combination of backing and rounding lowers vowel height and the combination of fronting and unrounding raises vowel height. Additional cross-linguistic analysis is needed not only to provide further support for these tentative patterns, but also to differentiate the relative contributions of rounding and backness to phonological shifts in vowel height.

### 2.4.3. Height Effects of Creaky Voice

This section is based on data reported in Den...
However, when the height of a back vowel was nearly equal to a lower front vowel, the highest front unrounded vowel was higher than the highest back vowel in 5 languages, the pattern /u/. Twenty-nine languages showed a height difference, with the relatively low back vowel /u/, but also /i u/, /i o/, and frequent absence of /u/ in vowel systems has been previous. Crothers, 1978; Disney.

2.4.3. Height Effects of Breathy and Creaky Voice
This section is based on the phonological data reported in Denning [1989], which examined evidence from 50 languages exhibiting an interaction between vowel height and phonation (taken to include phonation type, pitch, and voicing in adjacent consonants). In over half of these languages, the interaction involved the phonation types of breathy or creaky voice, sometimes co-occurring with other differences, such as pitch or tongue root position. For example, in Acoli, the breathy-nonbreathy distinction correlates with consistent differences in vowel height: breathy [i e ə ʊ] and nonbreathy [i e a o]. But the set of interacting factors may be quite complex, as in some dialects of Akan, where the vowels [i e ə ʊ] have breathy voicing, advanced tongue root position, and a lowered larynx, and [i e a o] have creaky voicing, normal tongue root position, and a raised larynx. Drawing on allophonic variation, phonological vowel inventories, historical change, and phonetic analyses, Denning [1989] concluded that, in languages with a correlation between vowel height and phonation: (a) greater vocal fold laxness (as in breathy voicing) is associated with higher vowels; (b) greater vocal fold tension (as in creaky voicing) is associated with lower vowels. It should be noted that the patterns in (a) and (b) are consistent with the predictions of the spectral integration model. The phonological patterns, however, would appear to be more general than expected in that the model makes no clear predictions concerning low vowels. It may be that, phonologically, low nonmodal vowels are less likely to shift height than nonlow ones. While the high and mid vowels in the languages cited by Denning consistently exhibit the predicted height effects of phonation type, the low vowels in several of the languages either fail to have a phonation contrast (as in the Akan example above) or show no height difference between modal and nonmodal vowels. Yet it is also the case that, in other languages, low vowels exhibit a clear effect of phonation type on vowel height. Noisy the frequent interaction of phonation type, voicing, and tongue root position in these languages, Denning [1989] suggested that the shifts in tongue body height might be the result of aerodynamic and neuromuscular factors, as well as biomechanical linkages involving tongue root retraction, laryngeal height, and vocal fold tension. It may well be that differences in the relative amplitudes of F₀ and F₁ interact with differences in F₁ frequency resulting from changes in articulatory configuration to enhance the perceptual effect of raising or lowering.

3. Discussion and Conclusion
This paper has focused on the interpretation of the spectral integration model in terms of its phonological consequences as a detailed illustration of one approach to the application of phonetic models to phonological systems. Using evidence from articulatory modeling and acoustic measures of naturally produced vowels, it was shown that manipulation of nasalization, oral cavity shape, and phonation type resulted in shifts in the low-frequency center of gravity of the vowel spectrum. Investigation of these spectral shifts in terms of the spectral integration model indicated that their perceptual consequence was to raise or lower
perceived vowel height, depending on the
direction of the center of gravity shift. But
while these changes in perceived height are
automatic consequences of (the model’s
characterization of) vowel processing, the
translation into phonological consequences
was shown to be mediated by both phonetic
and phonological considerations, yielding
restricted predictions concerning the pho-
notological effects of nasalization, backness,
rounding, breathiness and creak on vowel
height. Furthermore, even these restricted
predictions (taken to apply not to all lan-
guages, but only to those exhibiting a corre-
lation between vowel height and one or
more of these properties) must be viewed as
the output of a single component of the
speech mechanism, and hence as predic-
tions that could in principle be outweighed
by other forces.

However, despite these limitations, the
height shifts predicted by the model were
generally borne out by phonological pat-
terns of height differences between oral ver-
sus nasal vowels, back rounded versus front
unrounded vowels, and modal versus non-
modal vowels. (That certain other height ef-
fects not predicted by the model were also
found is not surprising since, as noted,
viewing the model in isolation ignores in-
fluences other than spectral integration.) The
overall parallel, then, both supports the val-
idity of the assumptions underlying the cur-
rent approach and leads to the conclusion
that the phonetic motivation for diverse
phonological processes of vowel raising and
lowering is, at least in part, the pheno-
menon of spectral integration.

Before considering what the conse-
quences of this conclusion might be for
phonological theory, it is noted that the lim-
itations on the translation of the model’s
output into phonological predictions are
not unique to the spectral integration
model, nor even to listener-based models in
general. Of the models presented in the lit-
erature, neither those interpreted as gener-
ating default settings [e.g., Lindblom,
or those viewed as imposing physical lim-
its [e.g., Ohala, 1981, 1983] derive excep-
tionless predictions for phonological sys-
tems. It would appear that only constraints
of the type ‘the human vocal mechanism
cannot produce the sound X’ or ‘the human
auditory system cannot differentiate be-
tween the sounds X and Y’ would yield
such predictions. Yet to the extent that such
constraints are known [see, e.g., Caffard,
1977], they fail considerably short of char-
acterizing the vowel or consonant space uti-
lized by the world’s languages [Lindblom,
1983, 1999; Ladefoged, 1985].

Of what value to phonological theory,
then, are phonological predictions of the
type generated by the spectral integration
model? Taking one of the goals of phono-
logical theory to be the delimitation of the
set of possible phonological systems, the
answer is that such predictions enable us, in
principle, to delimit the set of possible
phonetically motivated phonological pheno-
mena. The qualification ‘in principle’ is
needed because, in most cases, a given
model provides information concerning
only one component of the speech process.
So, for example, the spectral integration
model as interpreted here identifies only a
particular class of perceptually motivated
vowel height phenomena. In reality, then,
predictions that delimit the more general
set of ‘phonetically motivated’ phonolo-
gical phenomena will depend on the conver-
gence of the output of models of all speech
components. Fortunately, such conver-
gent predictions extend to components of the speech pro-
but there is increasing em-
teraction [Stevens, 1972,
1983; Lindblom and Mad

Implicit in this view of phonetic modeling to-
theory is the assumption that
be able to determine whe-
nomological phenomenon is
pecific factors. Inasmuch as
equivalent to saying that i
able to explain phonole
this assumption is taken
al. To quote Anderson [5]
still very much part of the
ologists to look for “s
ions” of phonological pho’nomic approach taken here, the
ions or predictions fall s
ive-nomological form rel
980] since the model can
ecessitating phono
vowel height. Rather, the
om model achieves what
ferred to as deductive-p
ations.

In concluding, it is
is there is considerable po-
thening the contribution
els to phonological theo-
ain limitations: inappro-
, others can be ad-
. One rectifiable lim-
adies of adequate descript-
verse phonetic and phone-
also, as noted, manip-
olated components of

Phonological predictions are to the spectral integration in listener-based models in the models presented in the literature those interpreted as general settings [e.g., Lindblom, Vestbury and Keating, 1986] or as imposing physical limits, 1981, 1983] derive exceptions for phonological systems that appear to only constraints the human vocal mechanism to the sound X or the human ear cannot differentiate between X and Y would yield the same. Yet to the extent that such a known [see, e.g., Cafton, 1985] considerably short of dialect-vowel or consonantal space universal's languages [Lindblom, 1985].

Due to phonological theory, phonological predictions of the spectral integration are one of the goals of phonological studies to be the delimitation of the phonological systems, the phonological predictions enable us to delimit the set of possible motivated phonological qualification 'in principle' is use, in most cases, a given level of information concerning the phonetics of the speech processes, the spectral integration model here identifies only a subset of perceptually motivated phenomena. In reality, then, at delimit the more generally motivated phonological model will depend on the convergence of models of all speech components. Fortunately, phonetics research would appear to be moving in the direction of such convergence: not only do phonetic studies extend to virtually all components of the speech process [Ohala, 1990], but there is increasing emphasis on their interactions [Stevens, 1972, 1989; Lindblom, 1983; Lindblom and Maddieson, 1988].

Implicit in this view of the contribution of phonetic modeling to phonological theory is the assumption that it is desirable to be able to determine whether or not a phonological phenomenon is motivated by phonetic factors. Inasmuch as this is essentially equivalent to saying that it is desirable to be able to explain phonological phenomena, this assumption is taken to be uncontroversial. To quote Anderson [1981, p. 497]: 'it is still very much part of the business of phonologists to look for phonetic explanations of phonological phenomena'. In the approach taken here, the offered explanations or predictions fall short of the deductive-nomological form required by Dinnissen [1980] since the model cannot be interpreted as necessitating phonological shifts in vowel height. Rather, the spectral integration model achieves what Ohala [1987] has referred to as deductive-probabilistic explanations.

In concluding, it is emphasized that there is considerable potential for strengthening the contribution of phonetic models to phonological theory. Although certain limitations are inherent in the approach, others can be addressed by further study. One rectifiable limitation is the scarcity of adequate descriptive accounts of diverse phonetic and phonological structures. Also, as noted, manipulation of models of isolated components of the speech mechanism needs to be supplemented with study of possible interactions. The output of different components may be in conflict [Lindblom, 1990] or may enhance each other [see, e.g., section 2.4.3 and Stevens et al., 1986] with respect to a particular property, but in both cases knowledge of their interaction should improve the 'fit' between the predictions derived from phonetic theory and observed phonological phenomena.

Finally, interpreting the output of phonetic models in terms of phonological consequences is hindered by lack of knowledge concerning the nature of certain phonetic-phonological interactions. Within listener-based models, for example, perhaps the strongest limitation on translating the perceptual effects predicted by the model into phonological consequences is evidence that, under certain conditions, listeners are able to adjust or compensate for otherwise expected perceptual consequences. But relatively little is known about the conditions which facilitate such compensation. In particular, little is known about the extent to which compensatory abilities are influenced by knowledge of a particular phonological system. Krakow et al. [1988], for example, suggested on the basis of judgments by American English listeners that, for vowels, such abilities depend in part on the phonological appropriateness of the context in which the vowel occurs, but cross-linguistic investigation is needed to support this position.

It is hoped, then, that further investigation will enhance our understanding of the conditions under which perceptual compensation does or does not take place and, more generally, shed light on the conditions under which a particular phonetic constraint becomes a stable phonological pattern. However, while further study might
enable us to better characterize these appropriate or prerequisite conditions, serious constraints on the predictive power of phonetic models will remain. Interpretation of the spectral integration model was restricted here to predictions concerning the direction of height shifts in languages exhibiting a correlation between vowel height and certain properties. Given that other phonetic and nonphonetic forces may counteract the predictions of an isolated model, this approach encounters exceptions (although inspection of exceptional languages should indicate, in some cases at least, the nature of the countervailing forces). Thus, for current phonetic models, predictive power may be limited to the claim that there is a greater-than-chance tendency for phonological systems to reflect a given model’s constraints. Yet the consequences of such limited predictions for phonological theory should not be minimized. Phonetic models cannot predict the structure of phonological systems, but they enable us to predict which phonological structures recur across languages and provide explanations for their occurrence.

Acknowledgments

This paper has benefited from the comments and other types of help, of many people, and I especially wish to thank Andrée Cooper, William Croft, Keith Denning, Sarah Hawkins, Rena Krakow, and Doug Whalen.

References


Bhat, D. N. S.: Two studies on nasalization; in Ferguson, Hyman, Ohala, Nasal and Nasality. Papers from a Symposium on Nasals and Nasalization, pp. 27–48 (Stanford University, Stanford 1975).


On Predicting Phonological Str...


Denning, K.: The diachronic c...

Dillen, D. A.: Phonological r...


Fascher-Jorgensen, E.: Phonetic (nurmed) vowels in Guaju... (Engin, Ankara 1989).


Fuks, H., Kawashima, T., T higher formants in the pe IEEE Trans. Audio Ele...

Goldstein, L.: Vowel shifts in... in Cohen, van... (Foris, Dordrecht 1983).

Hawkins, S., Stevens, K. N., A
On Predicting Phonological Structure


Lindblom, B.: Economy of speech gestures; in Mac-


On Predicting Phonological Structure


Prof. Patrick Speeter Beddor
Program in Linguistics
University of Michigan
1076 Frieze Building
Ann Arbor, MI 48109 (USA)