The strategy of generative phonol-
sler, Mareš, Phonologica 1972, pp.
nk, München 1975).

lfgang Dressler
prachwissenschaft
Vien

(Austria)

Phonetica 1991; 48: 149–179

Integrating Articulations in the Perception of Vowel Height

John Kingston

Department of Linguistics, South College, University of Massachusetts, Amherst, Mass., USA

Abstract. In vowels contrasting for height, a large number of articulations covary with tongue height, which is supposed to be the principal bearer of the contrast. However, attempts to link these covarying articulations to tongue movement physiologically have been largely unsuccessful, and the particular pattern of covariation appears to make more sense as a concerted effort to influence the perceived height of F_1. The experiments reported here used the Garner interference paradigm, modified to assess the perceptual primacy of stimulus dimensions, to show that the acoustic effects of two of these covarying articulations, velum height (nasalization) and rate of vocal fold vibration (pitch) are integrated perceptually with the acoustic effects of varying tongue height. This perceptual integration suggests that the different articulations are not independently perceived, contrary to the predictions of direct realist theories of speech perception, that articulatory events covary so as to enhance each other's perceptual effects, and that the surface phonological or initial phonetic representation of vowels might be quite richly specified (contrary to claims of phonetic underspecification).

1. Introduction

Many articulatory settings covary directly with the height of the tongue in vowels: the higher the vowel, the more advanced the tongue root [Lindau, 1975, 1978, 1979; Jackson, 1988], the higher the jaw [Lindblom and Sundberg, 1971], the more protruded the lips in back vowels and the more retracted the lips in front vowels [Linker, 1982], the higher the velum [Moll and Shriner, 1967; Lubker, 1968; Bell-Berti, 1976; Bell-Berti et al., 1979], and the higher the rate of vocal fold vibration [Peterson and Barney, 1952; House and Fairbanks, 1953; Lehiste and Peterson, 1961; Honda, 1987; Steele, 1986; Ohala and Eukel, 1987; Silverman, 1987]. Duration, on the other hand, varies inversely with tongue height [Lindblom, 1967; Lehiste, 1970; Westbury and Keating, 1980; Keating, 1985]. The conventional view is that these other articulatory settings are secondary to the height of the tongue in distinguishing one vowel from another, and may even be mechanical consequences of moving the tongue [for an argument of this type, see Ohala and Eukel, 1987]. Here, it will be argued that vowel 'height' is instead an integration of all these settings, with no priority, for articulatory control nor acoustic goals, given to the height of the tongue. The basis of this claim is that, aside from the duration difference, and the labial articulations, which affect F_2 and thus the percept of vowel backness most strongly, all these articulatory settings converge on a single effect, the manipulation of the perceived F_1. ('Perceived F_1' is used as shorthand in this paper for the listener's response to the distribution of energy in the lower part of the spectrum, which contains the first resonance, as well as the acoustic effects of a variety of other articulations.) It is to bring about this convergence that all these articulatory events covary as they do, and not some purely physiological dependency between them [see Diehl and Kluender, 1989, and Diehl, this volume, for similar views].

The available evidence indicates in fact that the articulatory settings that covary with tongue height are independently controlled. Many languages of East and West Africa exhibit contrasts for the position of the tongue root as well as for tongue height [Hall et al., 1974; Lindau, 1975, 1978, 1979; Jackson, 1988]; the velum is actively raised in higher vowels [Lubker, 1968]; F_0 is actively elevated through the contraction of the cricothyroid in higher vowels [Vilkman et al., 1989; Honda and Fujimura, 1989] and F_0 differs in the same direction between vowels of different heights in esophageal as well as laryngeal speech [Gandour and Weinberg, 1980; Fox and Trudeau, 1988]; and speakers differ in how much of the vertical position of the tongue is achieved by jaw as opposed to tongue raising or lowering [Ladefoged, 1990]. These demonstrations of independent control undermine any attempt to explain the observed covariation which yokes the other articulations physiologically to the movement of the tongue. Explanations of a different sort must be sought, like the one argued for here that unites these articulations in their effects on the perceived frequency of the lowest resonance of the vowel.

The convergence of the acoustic effects of these articulations could indicate that the distinctive features that represent vowel height contrasts, [high] and [low], or Clements's [1989] feature [open], do not, despite their names, refer simply to tongue height or degree of constriction, but are instead labels for acoustic or perceptual values [for similar views, see Ladefoged, 1980; Disner, 1983]. In a weak sense of this view, vowel height, i.e., perceived F_1, is always and primarily a function of tongue height, even though the other articulations that covary with tongue height enhance the perceived difference in F_1 of vowels produced with different tongue heights. In the strong sense of this view, on the other hand, vowel height contrasts do not necessarily imply differences in tongue height; instead the acoustic effects of each of the covarying articulations contribute equally to the perceived F_1. Height then is an abstract scale representing the covariation among these articulations jointly. Evidence from the behavior of vowels in phonological inventories, rules, and changes in support of this stronger view will be outlined in the last section of this paper.

If vowel height is an int
ulations in the phonetics,
of all these articulatory ev
acoustic effect also undern
of direct perception of
[Fowler, 1986, 1990], sin
argues that the articulat
properties in the acoustic s
individually recovered. Di
of speech gestures does n
possibility that during any
in the signal it may be un
an acoustic property is th
gesture that produces it is
rent segment or because o
with gestures of adjacent
ever, this theory does argue
guity will always be resolve
val is evaluated that is lon
veal any coarticulatory 's
acoustic properties [Fowl
1986; Krakow et al., 1988;
man, 1987]. However, none
ing articulations listed abov
latory origins, and if the
have the same acoustic effe
fect cannot be attributed, ex
to just one of them. That t
other than tongue height
acoustic properties of the vo
e.g., its pitch, nasality, etc.,
unless the listener can re
them perceptually (in any
height itself affects more th
tual separation would ind
acoustic effects of the articu
vary with tongue height m
different tongue heights ea
guish simply by adding to
between them. (Such a theo
course, provide any explan
these other articulations

' the vertical position of the
ieved by jaw as opposed to
g or lowering [Ladefoged,
demonstrations of independ-
ndermine any attempt to ex-
rved covariation which yokes
iculations physiologically to
: of the tongue. Explanations
sort must be sought, like the
or here that unites these arti-
heir effects on the perceived
the lowest resonance of the

rgence of the acoustic effects
ulations could indicate that
ve features that represent
contrasts, [high] and [low], or
989] feature [open], do not,
names, refer simply to tongue
gree of constriction, but are
for acoustic or perceptual
milar views, see Ladefoged,
, 1983]. In a weak sense of
wel height, i.e., perceived F_1,
nd primarily a function of
, even though the other artic-
covary with tongue height
perceived difference in F_1 of
uced with different tongue
e strong sense of this view,
hand, vowel height contrasts
ssarily imply differences in
; instead the acoustic effects
e covarying articulations con-
y to the perceived F_1. Height
ostract scale representing the
among these articulations
ence from the behavior of
onological inventories, rules,
in support of this stronger
outlined in the last section of

If vowel height is an integration of articulations in the phonetics, the convergence of all these articulatory events on a single acoustic effect also undermines the theory of direct perception of speech gestures [Fowler, 1986, 1990], since that theory argues that the articulatory sources of properties in the acoustic signal can all be individually recovered. Direct perception of speech gestures does not rule out the possibility that during any limited interval in the signal it may be uncertain whether an acoustic property is there because the gesture that produces it is part of the current segment or because of coarticulation with gestures of adjacent segments. However, this theory does argue that this ambiguity will always be resolved once an interval is evaluated that is long enough to reveal any coarticulatory source of such acoustic properties [Fowler and Smith, 1986; Krakow et al., 1988; see also Silverman, 1987]. However, none of the covarying articulations listed above have coarticulatory origins, and if these articulations have the same acoustic effect, then this effect cannot be attributed, except arbitrarily, to just one of them. That the articulations other than tongue height affect other acoustic properties of the vowel than its F_1, e.g., its pitch, nasality, etc., does not help unless the listener can reliably separate them perceptually (in any case, tongue height itself affects more than F_1). Perceptual separation would indicate that the acoustic effects of the articulations that covary with tongue height make vowels of different tongue heights easier to distinguish simply by adding to the differences between them. (Such a theory does not, of course, provide any explanation for why these other articulations covary consistently in the way they do.) The experiments reported in this paper are a first attempt to test whether the acoustic effects of these covarying articulations are instead integrated in the perception of vowel height, and the extent to which listeners can separate their articulatory origins.

Looking for this sort of integration is a test of a strong version of the auditory enhancement theory argued for by Diehl and his colleagues [Kluender et al., 1988; Diehl and Kluender, 1989; Diehl and Walsh, 1989, Diehl, this volume]. This theory claims that articulations covary in speech because the perceptual effect of one articulation exaggerates the perceptual value of another so that that speech sound becomes more distinct from one in which the two articulations covary in the opposite direction. For example, vowels are lengthened before [+voice] stop closures to exaggerate the perceived shortness of what is already a shorter closure than in [−voice] stops [Kluender et al., 1988]. The perceived duration of the [−voice] stop closure would of course be lengthened by the shorter vowel that precedes it. The perceptual effect of vowel duration on judgments of following stop closure duration is attributed to a general auditory mechanism Kluender et al. [1988] call 'durational contrast', and proponents of this theory argue that many cases of enhancement can be attributed to such mechanisms. This paper does not address the issue of whether enhancement of one perceptual effect by another arises out of such general auditory mechanisms, but instead the more general prediction of enhancement theory that the perceptual effects of covarying articulations merge into a single, more distinct perceptual object.

Table 1. Two-by-two stimulus array of the Garner paradigm

	Dimension 1	
	Low	High
Dimension 2		
High	A	B
Low	C	D

Dimension 1, e.g. Nasalization or Pitch; dimension 2, e.g. Tongue Height.

2. Methods and Stimuli

2.1. *What Kind of Experimental Design Is Needed?*

The hypothesis that all the articulations which covary with tongue height do so because their perceptual effects are the same suggests that listeners do not actually hear them as separable articulations, or at least not as articulations associated with any other contrast than vowel height. Accordingly, what is needed is an experimental design which allows a test of the extent to which the acoustic effects of these various articulations are perceptually integrated. More precisely, what is needed is a test of whether the perception of the value a vowel has along an acoustic dimension reflecting one articulation is influenced by its value along an acoustic dimension determined by another, covarying articulation. Without a demonstration that a vowel's perceptual value along one acoustic dimension depends on its acoustic value along another (and vice versa), then the only motivation for covarying articulations is that differences along more than one dimension at a time make vowels more distinct. Of interest here is the stronger claim that articulations covary because their perceptual effects exceed their sum and actually converge on a single, exaggerated perceptual value, for F_1.

At the same time, since each of these articulations participates in phonemic contrasts in its own right, their acoustic effects could also define orthogonal dimensions by which listeners sort vowels. For example, nasalization may make a low vowel sound lower, as a result of perceptual integration of the acoustic effects of nasalization with those of tongue height, and also simply nasalized, since that is an acoustic dimension orthogonal to vowel quality. To sum up, needed are a test of perceptual integration of orthogonal acoustic dimensions and a test of whether listeners can also classify stimuli along those dimensions individually. These two tests are provided by the Garner [1974] paradigm, adapted as suggested by Melara and Marks [1990], which is described in the next section.

2.2. *Integrability vs. Separability: The Garner Paradigm*

The integrability of apparently orthogonal stimulus dimensions has frequently been examined using a paradigm developed by Garner [1974]. In this paradigm, subjects are required to classify, as quickly as possible, stimuli selected from a two-by-two array defined by the dimensions whose integrability is being investigated (table 1 indicates the dimensions that were investigated in the experiments reported here).

The stimuli are supposed to be separated along each dimension so as to be equally discriminable. Stimuli are selected from this array to be presented to subjects for classification in a number of different ways (fig. 1):

(1) In *baseline* classification, stimuli are selected from just one side of the array and they thus differ along just one dimension: i.e. A vs. C or B vs. D on the Tongue Height dimension, and A vs. B or C vs. D on the Nasalization or Pitch dimensions.

(2) In *correlated* classification, stimuli are selected from opposite corners of the array, and they differ along both dimensions at once. Their values on these dimensions are either positively, B (both high) vs. C (both low), or negatively, A (low, high) vs. D (high, low), correlated.

(3) In *selective attention*, all four stimuli in the array are selected, but the perceiver must attend selectively to just one dimension while ignoring the other. In classification for Tongue Height, A and B are contrasted with C and D, while for Nasalization or Pitch, A and C are contrasted with B and D. The orthogonal difference along the other dimension is supposed to be irrelevant to the classification.

Fig. 1. Stimulus arrays at
(left) and 45° (right), illustrat
class membership in instances
the various tasks at the two ro
tions; the stimulus arrays at 22
would be intermediate. The ve
cal axis is (the inverse of) Ton
Height (F_1) and the horizontal a
is either Nasalization (N_2-N_1)
Pitch (F_0). Circles enclose stin
which belong to a class on a gi
block of trials, squares encl
stimuli that do not. At 0°, the ill
trated Baseline and Selective
tention tasks require judgments
Tongue Height; the Correla
task contrasts positively correla
stimuli; and the Divided Attent
task contrasts positively with ne
tively correlated stimuli.

(4) In *divided attention*, a
again selected, but the perceive
tion between the two dimension
two positively correlated stimu
guished from A and D, the tv
lated stimuli. Both dimensions
as is the nature of the correlatio
ues.

If the two dimensions of the
once, as they are in the correl
tion, and divided attention tas
mensions are integrated perce
separable, then performance
these tasks and the baseline ta
differ along just one dimensio

of perceptual integration of the
nasalization with those of tongue
imply nasalized, since that is an
ı orthogonal to vowel quality. To
:e a test of perceptual integration
ɔustic dimensions and a test of
can also classify stimuli along
individually. These two tests are
arner [1974] paradigm, adapted as
ra and Marks [1990], which is de-
t section.

ility vs. Separability:
Paradigm

ty of apparently orthogonal stim-
ıas frequently been examined us-
:veloped by Garner [1974]. In this
ts are required to classify, as
e, stimuli selected from a two-by-
by the dimensions whose integra-
:stigated (table 1 indicates the di-
'e investigated in the experiments

: supposed to be separated along
o as to be equally discriminable
:d from this array to be presented
ssification in a number of differ-

classification, stimuli are selected
: of the array and they thus differ
ıension: i.e. A vs. C or B vs. D on
t dimension, and A vs. B or C vs.
ıtion or Pitch dimensions.
'ed classification, stimuli are se-
ite corners of the array, and they
dimensions at once. Their values
ɔns are either positively, B (both
low), or negatively, A (low, high)
correlated.
: attention, all four stimuli in the
, but the perceiver must attend se-
ne dimension while ignoring the
ıtion for Tongue Height, A and B
h C and D, while for Nasalization
are contrasted with B and D. The
:nce along the other dimension is
relevant to the classification.

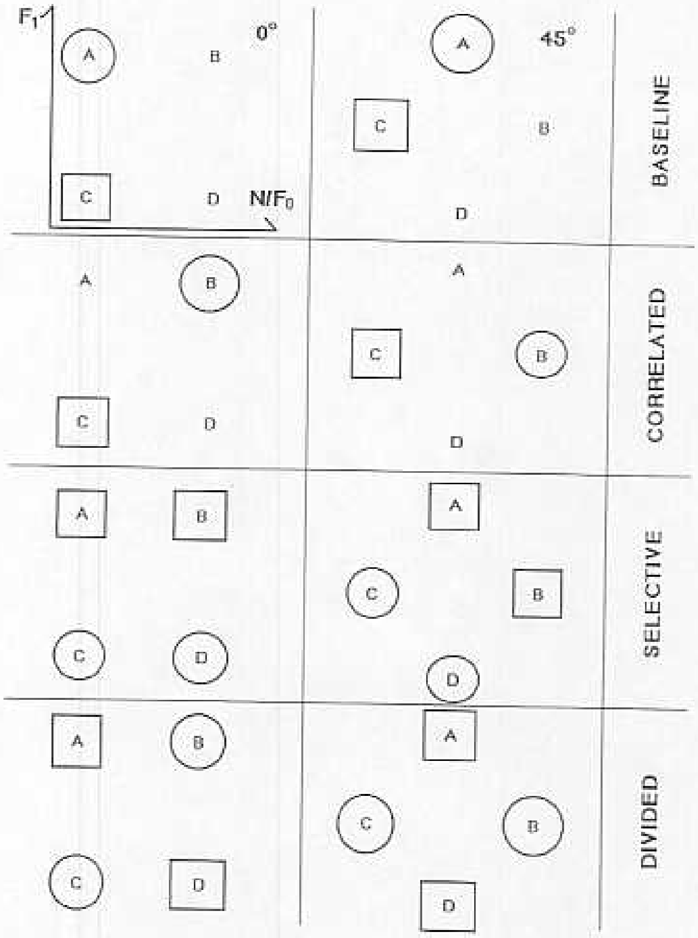

Fig. 1. Stimulus arrays at 0° (left) and 45° (right), illustrating class membership in instances of the various tasks at the two rotations; the stimulus arrays at 22.5° would be intermediate. The vertical axis is (the inverse of) Tongue Height (F_1) and the horizontal axis is either Nasalization (N_2-N_1) or Pitch (F_0). Circles enclose stimuli which belong to a class on a given block of trials, squares enclose stimuli that do not. At 0°, the illustrated Baseline and Selective Attention tasks require judgments of Tongue Height; the Correlated task contrasts positively correlated stimuli; and the Divided Attention task contrasts positively with negatively correlated stimuli.

(4) In *divided attention*, all four stimuli are again selected, but the perceiver must divide attention between the two dimensions, since B and C, the two positively correlated stimuli, have to be distinguished from A and D, the two negatively correlated stimuli. Both dimensions are always relevant, as is the nature of the correlation between their values.

If the two dimensions of the stimuli are varied at once, as they are in the correlated, selective attention, and divided attention tasks, and if these dimensions are integrated perceptually rather than separable, then performance will differ between these tasks and the baseline task, in which stimuli differ along just one dimension at a time. Integration of dimensions means that the value a stimulus has along one dimension changes its perceptual value along the other dimension. For example, a high value for Nasalization could make a vowel sound lower than its value for Tongue Height alone would lead one to expect.

Integration causes the stimuli to be classified more rapidly and accurately in the correlated than the baseline task, because the change in the stimuli's perceptual values along one dimension by their value along the other exaggerates the difference between the stimuli to be contrasted in the correlated task. If dimensions are perceptually separable, then performance might be facilitated because there are two rather than just one difference between the stimuli. Demonstrating integration therefore requires that performance also be worse on the selective attention task compared to baseline. With perceptually separable dimensions, variation on the unattended dimension would not affect accuracy in classifying the stimuli along the attended dimension in the selective attention task. Integration reduces performance in selective attention because the supposedly irrelevant differences along the unattended dimension change the stimuli's perceived value on the attended dimension. This increases the variability among the members of each class and with it the perceiver's uncertainty about whether a stimulus belongs to a class.

Having to attend to both dimensions at once would by itself reduce performance in the divided attention task, but integration should add to this difficulty. Integration should degrade performance even more in divided attention because the mutual changes in a stimulus's value along one dimension by its value along the other will make it more difficult to detect that the stimulus's values were positively or negatively correlated. With perceptually separable dimensions, on the other hand, the stimulus's value along each dimension can be independently measured and the polarity of the correlation between the values on the two dimensions would therefore be more easily determined.

The natural pattern of covariation between Nasalization or Pitch and Tongue Height could also influence performance in these classification tasks if listeners found the vowels in which the dimensions covaried as in natural speech more distinct than those in which the covariation went in the opposite direction. If this is so, then vowels in which Pitch and Tongue Height are positively correlated or in which Nasalization and Tongue Height are negatively correlated should be more easily distinguished (and less easily grouped together) than vowels in which these correlations have the opposite polarity. Thus performance should be better on the positively than negatively correlated task when vowels vary in Pitch and Tongue Height, but better for the negatively than positively correlated task when they vary in Nasalization and Tongue Height. Conversely, performance should be even poorer in the divided attention task since vowels which have been made more distinct by natural covariation will have to be classed together. Performance on the baseline and selective attention tasks should not be affected by any greater distinctiveness that may arise out of natural covariation of dimensions. In each of the baseline tasks, a distinction has to be drawn between a vowel which is more distinct because it exhibits natural covariation and one which is less distinct because it does not. In the selective attention tasks, each class always contains one more and one less distinct vowel. In both kinds of tasks, then, the effect of an additional distinctiveness that may arise out of natural covariation of dimensions is evenly distributed among the subtasks.

In summary, by altering the perceptual value of a stimulus along one dimension through its value along the other, integration should facilitate fast, accurate classification in the correlated tasks compared to the baseline task, but interfere with it in the selective and divided attention tasks. If integration also renders some of the stimuli more distinct than the others, perhaps those in which stimuli's values covary as they do naturally, then further effects will be observed in the correlated and divided attention tasks. With dimensions that are perceptually separable, neither facilitation nor interference is expected since the value a stimulus has along one dimension does not alter its perceptual value along the other.

2.3. Perceptual Primacy

2.3.1. Rotation of the Array

Melara and Marks [1990] argue that the perceptual integration of dimensions does not preclude independent evaluation of a stimulus along perceptually primary dimensions; these are the psychologically real dimensions in the stimulus space along which any stimulus is evaluated. They suggest instead that even if two dimensions are integrated in the perception of the *stimulus as a whole*, subjects may still extract the stimulus's *attributes*, i.e. its values along perceptually primary dimensions. Following Smith and Kemler [1978] and Grau and Kemler Nelson [1988], Melara and Marks [1990] show how rotating the array in the plane alters the perceptual interaction between its members and thereby reveals the orientation in the stimulus space corresponding to perceptually primary dimensions.

If the original orientation (0°) corresponds to perceptually primary but integrable dimensions, then rotation (fig. 1) increases the intraclass varia-

---

bility on the attended dimension i
distracting variability on the other
Rotation should thus cause mo
the selective attention task. Furthe
array is rotated 45°, stimuli which
be sorted into different classes, i.
vs. C, are actually identical on one
ing selective attention even more
On the other hand, rotation
sponses faster and more accurate i
tention task. As in the selective a
four stimuli from the array are pr
vided attention task, but the subje
tention between the two dimensio
positively correlated or congruent
from the negatively correlated or i
A and D. The facilitation that rota
in the divided attention task arises
ferences on the attended dim
orthogonal differences, and by a
bers of each class more closely w
primary perceptual dimensions. W
0 to 22.5°, the congruent stimuli,
differ more on the horizontal di
vertical, while for the incongruent
the vertical difference expands a
one shrinks. With further rotatio
gruent stimuli differ only along
mension and incongruent ones
vertical. As Melara and Marks [
this orientation the divided attenti
a same-different decision along
and should thus be faster and mo
the original orientation.
When dimensions are integra
ally primary, the difference in re
curacy between selective attent
should therefore increase with rot
lus axes since the stimulus array i
with these perceptually primary
difference between divided atte
response time or accuracy shoul
with rotation since attention can
on a single dimension.

2.3.2. Congruence

Melara and Marks [1990] des
sure of perceptual primacy, the
vs. incongruent variation along t
At issue is whether perceivers ca

ty greater distinctiveness that may
itural covariation of dimensions. In
seline tasks, a distinction has to be
a vowel which is more distinct be-
s natural covariation and one which
because it does not. In the selective
each class always contains one more
stinct vowel. In both kinds of tasks,
of an additional distinctiveness that
of natural covariation of dimensions
outed among the subtasks.
, by altering the perceptual value of
ig one dimension through its value
integration should facilitate fast, ac-
ation in the correlated tasks com-
eline task, but interfere with it in the
vided attention tasks. If integration
ne of the stimuli more distinct than
aps those in which stimuli's values
lo naturally, then further effects will
the correlated and divided attention
ensions that are perceptually separ-
ilitation nor interference is expected
a stimulus has along one dimension
s perceptual value along the other.

tual Primacy

on of the Array

Marks [1990] argue that the percep-
of dimensions does not preclude in-
iation of a stimulus along perceptu-
nensions; these are the psychologi-
isions in the stimulus space along
ulus is evaluated. They suggest in-
if two dimensions are integrated in
of the *stimulus as a whole*, subjects
the stimulus's *attributes*, i.e. its val-
otually primary dimensions. Follow-
Kemler [1978] and Grau and Kemler
Melara and Marks [1990] show how
y in the plane alters the perceptual
een its members and thereby reveals
n the stimulus space corresponding
primary dimensions.
al orientation (0°) corresponds to
imary but integrable dimensions,
g. 1) increases the intraclass varia-

---

bility on the attended dimension in addition to the distracting variability on the other dimension.

Rotation should thus cause more interference in the selective attention task. Furthermore, when the array is rotated 45°, stimuli which are supposed to be sorted into different classes, i.e. A vs. D and B vs. C, are actually identical on one dimension, making selective attention even more difficult.

On the other hand, rotation should make responses faster and more accurate in the divided attention task. As in the selective attention task, all four stimuli from the array are presented in the divided attention task, but the subject must divide attention between the two dimensions by sorting the positively correlated or congruent stimuli, B and C, from the negatively correlated or incongruent ones, A and D. The facilitation that rotation brings about in the divided attention task arises by increasing differences on the attended dimension, reducing orthogonal differences, and by aligning the members of each class more closely with the putatively primary perceptual dimensions. With rotation from 0 to 22.5°, the congruent stimuli, B and C, come to differ more on the horizontal dimension than the vertical, while for the incongruent stimuli, A and D, the vertical difference expands and the horizontal one shrinks. With further rotation to 45°, the congruent stimuli differ only along the horizontal dimension and incongruent ones differ only on the vertical. As Melara and Marks [1990] point out, at this orientation the divided attention task reduces to a same-different decision along either dimension and should thus be faster and more accurate than at the original orientation.

When dimensions are integrable but perceptually primary, the difference in response time or accuracy between selective attention and baseline should therefore increase with rotation of the stimulus axes since the stimulus array is no longer aligned with these perceptually primary dimensions. The difference between divided attention and baseline response time or accuracy should instead decrease with rotation since attention can focus increasingly on a single dimension.

2.3.2. Congruence

Melara and Marks [1990] describe another measure of perceptual primacy, the effect of congruent vs. incongruent variation along the two dimensions. At issue is whether perceivers can detect that two of the four stimuli in the array vary along the two dimensions in the same direction – both high or both low –, while the other two vary in opposite directions – one high and one low. The perceiver's ability to detect congruence or its absence depends on a prior ability to detect whether a stimulus has a high or low value along the two dimensions individually, which should only exist if those dimensions are perceptually primary. If stimulus attributes are not classified in terms of these individual dimensions, then it would not be possible to tell whether a stimulus had high or low values along both dimensions. The dimensions would not in fact exist perceptually as measures of stimulus value, and since no determination of congruence could be made, no difference in performance should be observed between the putatively congruent and incongruent stimuli.

2.4. The Stimuli

The perceptual integrability of the acoustic effects of two articulations that covary with tongue height, velum height and rate of vocal fold vibration – referred to here as Nasalization and Pitch – as well as their perceptual primacy, are examined in experiments designed on the models of Garner [1974] and Melara and Marks [1990].

The effect on formant frequencies of different degrees of tongue height, lip, and larynx position were determined through the use of a software model, Vocal 2.1 [kindly provided by Peter Ladefoged], which allows simulation of vocal tract shapes, transforms them to area functions, and then calculates the formant frequencies such a shape would produce [Ladefoged, n.d.; see also Lindau, 1975, 1978, 1979; Jackson, 1988]. (Henceforth, Tongue Height will refer simply to the position of the tongue in the mouth, or to its principal acoustic correlate, F_1. Vowel height, on the other hand, will be reserved to refer to the phonological contrast between vowels, and it should not be taken to grant any priority to tongue height or position in conveying these contrasts.) Because this work will later be extended to the effects of lip rounding on perceived vowel height, these investigations focused on vocal tract shapes appropriate to back vowels. In designing the stimuli used here, the effects of Tongue Height differences were simulated across a 2-mm

Table 2. First three formant frequencies (in Hz) resulting from simulations of three degrees of Tongue Height

	F_1	F_2	F_3
+1 mm	385	1,030	2,360
0 mm	430	1,080	2,320
−1 mm	450	1,070	2,320

range from 1 mm below the neutral position to 1 mm above it, with the further assumption of 2 mm of lip protrusion, 2 mm of larynx depression, and a lip aperture 13 mm wide [see Lindblom and Sundberg, 1971, Riordan, 1977, and Linker, 1982, for discussion of these articulations and their acoustic effects.] With each millimeter of tongue raising, the tongue was advanced a third of a millimeter to simulate the more forward position of higher back vowels. The resulting values (rounded to the nearest 5 Hz) for the first three formants are given in table 2.

Although Tongue Height clearly affects F_2 and F_3 as well as F_1, for simplicity's sake it was decided to employ constant values for these two formants in the stimuli; the values used were 1,050 and 2,350 Hz, respectively. Variation in F_1 alone of the 65 Hz that 2 mm difference in Tongue Height produces is more than sufficient to produce a different vowel quality. These formant frequencies were then used to synthesize vowels through the cascade branch of a Klatt synthesizer. Nasalization was synthesized by placing a nasal zero (N_z) between a nasal (N_1) and oral pole (F_1) (the former was always 150 Hz lower than the latter); the further the nasal zero is from either of the two flanking poles, the more nasalized the resulting vowel sounds [Stevens et al., 1987]. Although nasalized vowels are typically weaker than oral ones [House and Stevens, 1956], the peak amplitude of all vowels were normalized to within 0.2 dB of one another, which would largely prevent listeners from using an amplitude difference in the Nasalization by Tongue Height experiment. This was done so that performance would be determined by the effects on the spectrum of separating the nasal zero from the nasal pole by various amounts. Any reduction in intensity in the region of the first formant in the nasalized vowels is of course not affected by this normalization, so the lower intensity in that region of the spectrum which is found in natural nasalized vowels [House and Stevens, 1956] could have been used by the listeners in the Nasalization by Tongue Height experiment. Finally, F_0 was simply varied directly.

Pilot experiments in which listeners discriminated between members of stimulus pairs differing by varying amounts along the dimensions of Tongue Height (the inverse of F_1), Nasalization and Pitch were then run to determine rough just noticeable differences along these various dimensions. For F_1, a difference of about 25 Hz was reliably detected; for Nasalization, an N_z–N_1 difference of 10 Hz was distinguishable from one of 50 Hz; and for Pitch, just a 5 Hz difference was detectable. Since differences are easiest to detect when two stimuli are presented side by side for comparison, and the stimuli were to be presented individually in the experiments reported here, somewhat larger differences than the minimum ones detectable were employed in constructing the stimulus arrays.

Finally, it is important to note that varying Nasalization and Pitch does not directly affect the measured frequency of F_1 in these stimuli, only Tongue Height differences do. However, it is predicted that the acoustic effects of these articulations will affect the perceived frequency of F_1. It is this prediction in fact that these experiments are designed to test.

The stimuli for the two sets of arrays, at the various rotations, are given in table 3. Henceforth in reference to these arrays, a distinction will be drawn strictly between Tongue Height which increases top to bottom and F_1 which increases from bottom to top. Nasalization and Pitch increase from left to right (fig. 1).

The differences in F_1, the inverse of Tongue Height, are the same for the two experiments. F_0 was set to the intermediate value of 125 Hz in all the stimuli in the Nasalization by Tongue Height experiment, while the difference between the nasal formant and the nasal zero was set to 0 in the Pitch by Tongue Height experiment, eliminating any nasalization.

In terms of the acoustic measures themselves, F_1, N_z–N_1, and F_0, the congruent stimuli in both sets of arrays would be B and C since it is these stimuli in which values along each pair of acoustic dimensions are either both high or both low, while

---

Table 3. Stimuli for Nasalization

Nasalization *by Tongue Height*	F_1 N_z N_1 N_z-N_1
	F_1 N_z N_1 N_z-N_1
Pitch *by Tongue Height*	F_1 F_0
	F_1 F_0

in A and D they have opposite valu
two arrays differ in which pair of
values in a way best corresponding
height contrasts. If more Nasal
vowel sound lower for a given To
inverse of F_1) and vice versa, then
els are the most distinct in vowel h
salization by Tongue Height arrays
is in accord with acoustic cong
higher Pitch makes a vowel of
Height sound higher, then A and D
tinct in vowel height for the Pitch b
arrays, and distinctness does not a
congruence. For the sake of consis
pretation of the arrays, B and C w
the congruent stimuli and A an
ones in all future discussion.

Since error rate rather than rea
measure of performance in these
vowels were made quite brief, l

/ this normalization, so the lower in-
:egion of the spectrum which is found
:alized vowels [House and Stevens,
.ve been used by the listeners in the
y Tongue Height experiment. Finally
varied directly.
iments in which listeners discrimi-
members of stimulus pairs differing
mounts along the dimensions of
t (the inverse of F_1). Nasalization,
then run to determine rough just no-
nces along these various dimensions.
:ence of about 25 Hz was reliably de-
salization, an N_z-N_1 difference of
inguishable from one of 50 Hz; and
a 5 Hz difference was detectable.
:es are easiest to detect when two
sented side by side for comparison,
were to be presented individually in
s reported here, somewhat larger dif-
the minimum ones detectable were
nstructing the stimulus arrays.
; important to note that varying Na-
Pitch does not directly affect the
uency of F_1 in these stimuli, only
differences do. However, it is pre-
acoustic effects of these articulations
perceived frequency of F_1. It is this
'act that these experiments are de-

for the two sets of arrays, at the vari-
re given in table 3. Henceforth in ref-
arrays, a distinction will be drawn
Tongue Height which increases top
F_1 which increases from bottom to
on and Pitch increase from left to

nces in F_1, the inverse of Tongue
same for the two experiments. F_0
termediate value of 125 Hz in all the
asalization by Tongue Height exper-
e difference between the nasal for-
asal zero was set to 0 in the Pitch by
experiment, eliminating any nasali-

the acoustic measures themselves,
l F_0, the congruent stimuli in both
vould be B and C since it is these
ı values along each pair of acoustic
either both high or both low, while

Table 3. Stimuli for Nasalization by Height and Pitch by Height arrays

		Rotation					
		0°		22.5°		45°	
		A	B	A	B	A	B
Nasalization by Tongue Height	F_1	460	460	465	453	466	445
	N_z	320	370	337	371	351	365
	N_1	310	310	315	303	316	295
	N_z-N_1	10	60	22	68	35	70
		C	D	C	D	C	D
	F_1	430	430	437	425	445	424
	N_z	290	340	289	324	295	309
	N_1	280	280	287	275	295	274
	N_z-N_1	10	60	2	49	0	35
		A	B	A	B	A	B
Pitch by Tongue Height	F_1	460	460	465	453	466	445
	F_0	120	130	122	131	125	132
		C	D	C	D	C	D
	F_1	430	430	437	425	445	424
	F_0	120	130	119	128	118	125

in A and D they have opposite values. However, the two arrays differ in which pair of stimuli combine values in a way best corresponding to natural vowel height contrasts. If more Nasalization makes a vowel sound lower for a given Tongue Height (the inverse of F_1) and vice versa, then the B and C vowels are the most distinct in vowel height for the Nasalization by Tongue Height arrays, and distinctness is in accord with acoustic congruence, but if a higher Pitch makes a vowel of a given Tongue Height sound higher, then A and D are the most distinct in vowel height for the Pitch by Tongue Height arrays, and distinctness does not agree with acoustic congruence. For the sake of consistency in the interpretation of the arrays, B and C will be considered the congruent stimuli and A and D incongruent ones in all future discussion.

Since error rate rather than reaction time was the measure of performance in these experiments, the vowels were made quite brief, lasting only 65 ms from the onset of energy to its offset, in order to reduce accuracy overall. Amplitude rose steeply from 0 to 57 dB in the first 10 ms and then more gradually to 60 dB in the next 10 ms, remained at 60 dB for 25 ms, and then fell back to 0 dB in mirror image of its rise (recall that the peak amplitude in the output waveform was also normalized for all stimuli). The frequencies of the first oral and nasal formants, the nasal zero, and the fundamental were constant throughout each stimulus, and the frequencies of formants above the first as well as all formant bandwidths were constant across stimuli.

2.5. Connections among Design Components

If integration of dimensions occurs, performance on the correlated task should be faster and more accurate than that on the baseline tasks, while performance of the selective and divided attention tasks should be slower and less accurate. If integrations shows up in these experiments, it would show that the perception of vowel quality integrates differences in the acoustic signal due to velar or laryngeal articulations with those due to movements of the tongue.

Even if integration occurs, listeners may still sort the stimuli by their values along the individual dimensions. For example, a vowel may still be perceived as having a high Tongue Height when Nasalization is also high, even though its perceived Tongue Height is lower than it would be if Nasalization were low. Conversely, a vowel may still be perceived as having a low degree of Nasalization when Tongue Height is high, even though its perceived Nasalization is higher than it would be if Tongue Height were low. If listeners can still judge a stimulus's value on these dimensions individually, even if that value is distorted by the stimulus's value on the other dimension, then the dimensions are perceptually primary, in accord with their traditional labelling with different features, [high] and [low] (Tongue Height), [nasal], and [High] (Pitch). As a result of integration, the listener's percept of the vowel's height is more than just an additive composite of its values for each of these dimensions, but at the same time the vowel is recognizable as being in large part composed of settings along each of these dimensions, because they are also perceptually primary.

Perceptual primacy is assessed in two ways. First, performance on selective and divided attention compared to the baseline tasks should change in opposite ways with rotation of the stimulus array away from its orientation to the original stimulus dimensions. If that orientation aligns with perceptually primary dimensions, then performance on a selective attention task should decline, since intraclass variability is increased with rotation. Performance on the divided attention task should, on the other hand, improve, since rotation reduces the intraclass variability with respect to the primary dimensions. Second, with perceptually primary dimensions, congruent stimuli should be perceived more reliably than incongruent ones, since in the congruent stimuli both stimulus dimensions vary in the same direction, while in the incongruent ones, they vary in opposite directions. If it should turn out that vowels which exhibit natural covariation for these dimensions are more distinct from one another than those which do not, then performance on tasks in which the INcongruent stimuli must be distinguished should actually be better in the Pitch by Tongue Height experiment (in the Nasalization by Tongue Height experiment acoustic congruence coincides with the natural pattern of covariation of those dimensions).

2.6. Method of Presentation

The stimuli were presented to subjects seated in a sound-treated room, at comfortable levels over headphones to one ear, the right for all subjects but 1, who claimed his left ear was better. The subjects ran the experiment themselves from a video display terminal through which they entered their responses and controlled their progress from block to block. Each block of trials began with a series of practice trials in which the stimuli that belonged to the category were presented in alternation with those that did not. Whether a stimulus belonged to the category was indicated by presenting the words 'yes' or 'no' on the computer screen immediately before the subject heard the stimulus. The message remained on the screen for 500 ms, and then the subject heard the stimulus. They were then to respond by striking labelled 'yes' and 'no' keys (these keys were the '1' and '3' keys of the keypad and the subjects struck them with the index fingers of their left and right hands, respectively). The next trial began 250 ms after their response. After the subjects had heard 8 repetitions of each stimulus in training, the test trials began. In each block of test trials, they heard 24 instances of each stimulus, in random order. Pilot work had shown that to prevent subjects from responding at chance on the more difficult tasks, they needed to be reminded of what sounds belonged to the category during the test trials; accordingly, 8 of the 24 test trials were preceded by 'yes' or 'no', as in the practice trials, while the remaining 16 were preceded by 'guess' (which trials were preceded by 'yes', 'no' or 'guess' was also random). Only their performance on the 16 trials where they had to guess was included in the results. The subjects were instructed to respond as quickly as possible to both the practice and test trials, and if anything to sacrifice accuracy to speed. They were also told to give

---

their immediate impression of whe
belonged to the category and not
when the decision was difficult.

There were two sessions, one fc
tion by Tongue Height arrays and t
Pitch by Tongue Heights arrays. H
did Nasalization by Tongue Heigl
other half Pitch by Tongue Height fi
session, there were 27 blocks of tria
9 tasks by 3 rotations. The 9 tasks
tasks (A vs. B, C vs. D, A vs. C,
2 correlated tasks (B vs. C and A vs
attention tasks (A and B vs. C and
vs. B and D), and 1 divided attentio
vs. A and D), and the 3 rotations v
45°. Subjects heard all 9 tasks with
gether, but the order in which they
tion was determined by a Latin sq
order in which the 9 tasks were pr
rotation, expect that the first task tl
to perform on any rotation was a ba
restriction ensured that they started
finding out what one of its 'dime
nally, whether a given stimulus or
to the category on a given task was
across subjects.

2.7. Subjects and the Pract of the Experiments

Subjects in this experiment we
volunteers from the Cornell Univ
who answered posted advertisemer
dents, 5 undergraduate and 3 gra
had some background in linguisti
beginning their 2nd year of grad
guistics, but all were entirely ign
pose of the experiment. Six of t
women and the other 2 men. Each
in both sessions of the experimen
took approximately an hour and
1 subject, the two sessions were
days; for that subject the first sess
morning and the second in the
same day. Subjects paced themse
and were told that they could tak
they wished between blocks.

: distinct from one another than those
then performance on tasks in which
ent stimuli must be distinguished
y be better in the Pitch by Tongue
nent (in the Nasalization by Tongue
ment acoustic congruence coincides
al pattern of covariation of those di-

od of Presentation

i were presented to subjects seated in
d room, at comfortable levels over
one ear, the right for all subjects but
d his left ear was better. The subjects
nent themselves from a video display
gh which they entered their responses
their progress from block to block.
trials began with a series of practice
the stimuli that belonged to the cate-
sented in alternation with those that
her a stimulus belonged to the cate-
ated by presenting the words 'yes' or
nputer screen immediately before the
the stimulus. The message remained
or 500 ms, and then the subject heard
hey were then to respond by striking
nd 'no' keys (these keys were the '1'
the keypad and the subjects struck
index fingers of their left and right
ively). The next trial began 250 ms
ponse. After the subjects had heard
each stimulus in training, the test
each block of test trials, they heard
each stimulus, in random order. Fi-
own that to prevent subjects from re-
ance on the more difficult tasks, they
minded of what sounds belonged to
ring the test trials; accordingly, 8 of
were preceded by 'yes' or 'no', as in
ls, while the remaining 16 were pre-
ss' (which trials were preceded by
uess' was also random). Only their
n the 16 trials where they had to
ded in the results. The subjects were
spond as quickly as possible to both
d test trials, and if anything to sacri-
) speed. They were also told to give

---

their immediate impression of whether a stimulus belonged to the category and not to reflect even when the decision was difficult.

There were two sessions, one for the Nasalization by Tongue Height arrays and the other for the Pitch by Tongue Heights arrays. Half the subjects did Nasalization by Tongue Height first and the other half Pitch by Tongue Height first. Within each session, there were 27 blocks of trials, consisting of 9 tasks by 3 rotations. The 9 tasks were 4 baseline tasks (A vs. B, C vs. D, A vs. C, and B vs. D), 2 correlated tasks (B vs. C and A vs. D), 2 selective attention tasks (A and B vs. C and D and A and C vs. B and D), and 1 divided attention task (B and C vs. A and D), and the 3 rotations were 0, 22.5, and 45°. Subjects heard all 9 tasks within a rotation together, but the order in which they heard each rotation was determined by a Latin square, as was the order in which the 9 tasks were presented within a rotation, expect that the first task that a subject had to perform on any rotation was a baseline task. This restriction ensured that they started each rotation by finding out what one of its 'dimensions' was. Finally, whether a given stimulus or stimuli belonged to the category on a given task was counterbalanced across subjects.

2.7. Subjects and the Practical Conditions of the Experiments

Subjects in this experiment were 8 paid ($5/h) volunteers from the Cornell University community who answered posted advertisements; all 8 were students, 5 undergraduate and 3 graduate. Five of 8 had some background in linguistics, in fact 2 were beginning their 2nd year of graduate study of linguistics, but all were entirely ignorant of the purpose of the experiment. Six of the subjects were women and the other 2 men. Each subject took part in both sessions of the experiment, each of which took approximately an hour and a half. For all but 1 subject, the two sessions were run on different days; for that subject the first session was run in the morning and the second in the afternoon of the same day. Subjects paced themselves in a session, and were told that they could take breaks whenever they wished between blocks.

Since each stimulus was presented 16 times and there were 8 subjects, the results are based on 128 responses to each stimulus in each of the 9 tasks at each rotation. Given that there were 16 repetitions, of 4 stimuli, in 9 tasks, at 3 rotations, in 2 experiments, each subject produced 3,456 responses.

3. Results

3.1. Introduction

All subjects reported somehow that vowel quality was manipulated in both experiments. All also noticed that Pitch was manipulated in the Pitch by Tongue Height experiment, but the only subjects who described the orthogonal variable as Nasalization in the other experiment either had some linguistic training or experience with a language such as French in which nasalization is contrastive on vowels. The other subjects described some of the vowels in the Nasalization by Tongue Height experiment as muffled or otherwise obscured. These reports suggest that the subjects were able to detect the principal acoustic manipulations of these experiments, but they do not require that Nasalization, Pitch, and Tongue Height be perceptually primary dimensions. All the subjects also reported far more difficulty with the selective and divided attention blocks in which two stimuli belonged to each class rather than just one. This difficulty could reflect interference from perceptual integration, or just the greater difficulty in forming a class of two rather than just one member. Both of these ambiguities are resolved by the analysis of the subjects' performance.

The dependent measure in these experiments was error rate as a proportion of the

Table 4. Mean error rates (with standard errors), as a proportion of total responses for each combinatio of Task Type and Rotation: Nasalization by Tongue Height and Pitch by Tongue Height

Task	Rotation							
	0°		22.5°		45°		mean	
	mean	SE	mean	SE	mean	SE	mean	SE
Nasalization by Tongue Height								
Baseline	0.148	0.018	0.160	0.029	0.151	0.037	0.153	0.016
Negative correlation	0.008	0.005	0.020	0.012	0.055	0.021	0.027	0.009
Positive correlation	0.090	0.042	0.215	0.071	0.156	0.067	0.154	0.035
Selective	0.190	0.022	0.287	0.051	0.282	0.036	0.253	0.023
Divided	0.441	0.027	0.498	0.030	0.379	0.037	0.440	0.020
mean	0.175	0.026	0.236	0.031	0.205	0.026	0.205	0.016
Pitch by Tongue Height								
Baseline	0.074	0.017	0.057	0.018	0.064	0.028	0.065	0.012
Negative correlation	0.035	0.015	0.020	0.008	0.031	0.019	0.029	0.008
Positive correlation	0.027	0.007	0.039	0.018	0.023	0.010	0.030	0.007
Selective	0.191	0.034	0.235	0.027	0.278	0.017	0.235	0.016
Divided	0.339	0.049	0.289	0.057	0.339	0.050	0.322	0.029
mean	0.134	0.023	0.128	0.022	0.147	0.025	0.136	0.013

total responses for that subject, task, and rotation. Preliminary analyses of variance (ANOVA) yielded no significant interactions between Subject and the other independent variables, so the requirements for a randomized complete block design were met here.

3.2. *Assessing Perceptual Integration: Overall Error Rates*

Mean error rates (with standard errors) obtained in the two experiments for each combination of Task Type and Rotation are presented in table 4 (positive correlation refers to the task where the stimuli either had high or low values on both acoustic dimensions, B and C, while negative correlation refers to the task where the stimuli had a high value on one acoustic dimension and a low value on the other, A and D).

Differences in overall error rate were assessed in terms of a three-way ANOVA with Experiment (Nasalization or Pitch), Rotation (0, 22.5, and 45°), and Task Type (baseline, negatively correlated, positively correlated, selective attention, or divided attention) as independent variables. The main effect of differences in Experiment was quite significant, $F_{(1, 203)} = 34.987$, $p < 0.0001$, with substantially more errors overall (51%) in the Nasalization by Tongue Height than the Pitch by Tongue

---

Fig. 2. Mean error rates (with standard errors), expressed as the proportion of the responses in a given task for each subject: the effects of Rotation in the Nasalization vs. Pitch by Tongue Height experiments (**a**) and the effects of Task Type in the Nasalization vs. Pitch by Tongue Height experiments (**b**).

Height experiment. Rotation
icant $F_{(2, 203)} = 2.000$, p = 0.1
was highly significant, $F_{(3,}$
p <0.0001, and post hoc tes
showed that all pairwise cc
tween Task Types, except th
positively correlated and 1
were significant at 0.05: ove
were significantly lower
correlated than baseline tas
cantly higher for selective at
higher for divided attention

' total responses for each combination
h by Tongue Height

45 °		mean	
mean	SE	mean	SE
0.151	0.037	0.153	0.016
0.055	0.021	0.027	0.009
0.156	0.067	0.154	0.035
0.282	0.036	0.253	0.023
0.379	0.037	0.440	0.020
0.205	0.026	0.205	0.016
0.064	0.028	0.065	0.012
0.031	0.019	0.029	0.008
0.023	0.010	0.030	0.007
0.278	0.017	0.235	0.016
0.339	0.050	0.322	0.029
0.147	0.025	0.136	0.013

low values on both acoustic di-
and C, while negative correla-
the task where the stimuli had
on one acoustic dimension and
on the other, A and D).

es in overall error rate were as-
rms of a three-way ANOVA
ment (Nasalization or Pitch),
22.5, and 45 °), and Task Type
gatively correlated, positively
elective attention, or divided
independent variables. The
of differences in Experiment
significant, $F_{(1,203)}$ = 34.987,
with substantially more errors
%) in the Nasalization by
ght than the Pitch by Tongue

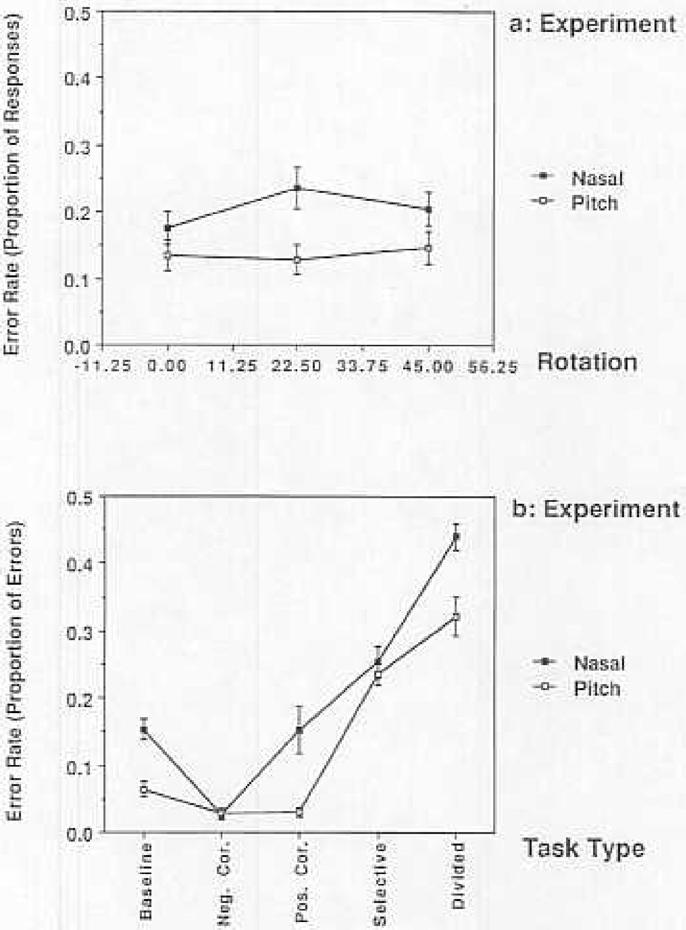

Fig. 2. Mean error rates (with standard errors), expressed as the proportion of the responses in a given task for each subject: the effects of Rotation in the Nasalization vs. Pitch by Tongue Height experiments (a) and the effects of Task Type in the Nasalization vs. Pitch by Tongue Height experiments (b).

Height experiment. Rotation was not significant $F_{(2,203)} = 2.000$, $p = 0.1379$. Task Type was highly significant, $F_{(3,203)} = 116.591$, $p < 0.0001$, and post hoc tests (Scheffé's S) showed that all pairwise comparisons between Task Types, except that between the positively correlated and baseline tasks, were significant at 0.05: overall error rates were significantly lower for negatively correlated than baseline tasks, and significantly higher for selective attention and yet higher for divided attention tasks.

Interpreting the different effects of negative vs. positive correlation requires looking at the interactions in this analysis, one of which reached significance, and another which approached it. The interaction between Experiment and Rotation, $F_{(2,203)} = 2.887$, $p = 0.058$ (fig. 2a), was almost significant because rotation affects error rates more in the Nasalization by Tongue Height than Pitch by Tongue Height experiment, and because more errors occurred at 22.5 ° than at other rotations. The differences be-

Table 5. Mean differences (with standard errors) between selective or divided attention and baselin error rates by Rotation: Nasalization by Tongue Height and Pitch by Tongue Height

Task	Rotation							
	0 °		22.5 °		45 °		mean	
	mean	SE	mean	SE	mean	SE	mean	SE
Nasalization by Tongue Height								
Selective – Baseline	0.042	0.022	0.127	0.038	0.131	0.047	0.100	0.022
Divided – Baseline	0.293	0.029	0.338	0.051	0.228	0.032	0.286	0.023
mean	0.168	0.037	0.232	0.041	0.179	0.030	0.193	0.021
Pitch by Tongue Height								
Selective – Baseline	0.118	0.030	0.179	0.023	0.215	0.019	0.170	0.016
Divided – Baseline	0.266	0.049	0.232	0.043	0.274	0.054	0.257	0.027
mean	0.191	0.034	0.201	0.024	0.245	0.029	0.214	0.017

tween Experiments in the effects of Task Type reflected in the second significant interaction, $F_{(4, 203)} = 4.818$, $p = 0.001$ (fig. 2b), include: a drop in error rate for both the positively and negatively correlated tasks compared to baseline performance in the Pitch by Tongue Height experiment, but only for the negatively correlated task in the Nasal by Tongue Height experiment vs. substantial increases in error rate for both experiments in the selective and divided attention tasks. The interference indicated by the poor performance on these last two tasks indicates integration of both Nasalization and Pitch with Tongue Height, but integration also predicts facilitation on both the positively and negatively correlated tasks. The absence of such facilitation in the positively correlated task in the Nasalization by Tongue Height experiment thus requires explanation. This difference between the two experiments in the behavior of positively vs. negatively correlated stimuli is shown to arise from differences in perceptual primacy between Nasalization and Pitch relative to Tongue Height in the evaluation of the effects of congruence in section 3.4.

3.3. Assessing Perceptual Primacy

Perceptual primacy of the dimensions of Tongue Height, Nasalization, or Pitch was assessed first by looking at whether the difference between selective attention and baseline performance gets larger with rotation from 0 to 45 °, while the difference between divided attention and baseline performance gets smaller. If Rotation interacts with Task Type in this way, then the dimensions corresponding to the original orientation are perceptually primary. More generally, the interaction of rotation with these

Fig. 3. Difference in mear error rates (with standard error for [Selective Attention – Baseline and [Divided Attention – Baseline] pooled across Nasalization anc Pitch by Tongue Height experi ments (a); Nasalization by Tongu Height alone (b) vs. Pitch b Tongue Height alone (c).

two kinds of differences allc tually primary dimensions to they are those dimensions a rotation at which the selecti baseline differences are sn divided attention vs. basel largest.

These two kinds of differ culated for each task, rotati and serve as the dependen three-way ANOVA in whic dent variables are Rotatio

ve or divided attention and baseline
Tongue Height

45°		mean	
mean	SE	mean	SE
0.131	0.047	0.100	0.022
0.228	0.032	0.286	0.023
0.179	0.030	0.193	0.021
0.215	0.019	0.170	0.016
0.274	0.054	0.257	0.027
0.245	0.029	0.214	0.017

vs. negatively correlated stim-
o arise from differences in per-
acy between Nasalization and
to Tongue Height in the eval-
: effects of congruence in sec-

ing Perceptual Primacy

primacy of the dimensions of
ht, Nasalization, or Pitch was
by looking at whether the dif-
een selective attention and
ormance gets larger with rota-
45°, while the difference be-
d attention and baseline per-
s smaller. If Rotation interacts
e in this way, then the dimen-
onding to the original orienta-
eptually primary. More gener-
raction of rotation with these

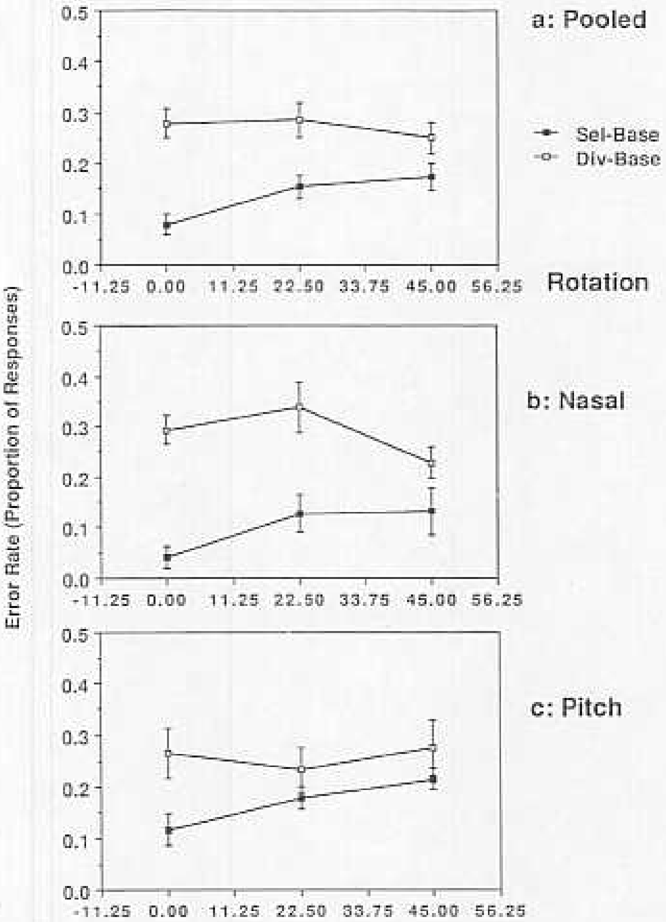

Fig. 3. Difference in mean error rates (with standard error) for [Selective Attention – Baseline] and [Divided Attention – Baseline]: pooled across Nasalization and Pitch by Tongue Height experiments (**a**); Nasalization by Tongue Height alone (**b**) vs. Pitch by Tongue Height alone (**c**).

two kinds of differences allows the perceptually primary dimensions to be discovered; they are those dimensions aligned with the rotation at which the selective attention vs. baseline differences are smallest and the divided attention vs. baseline differences largest.

These two kinds of differences were calculated for each task, rotation, and subject and serve as the dependent measure in a three-way ANOVA in which the independent variables are Rotation, Experiment, and Attention (= subset of Task Types; table 5).

Neither Rotation nor Experiment was significant, though Attention was, $F_{(1,77)} = 44.321$, $p < 0.0001$, reflecting the poorer perfomance overall in divided compared to selective attention. The interaction of Experiment by Attention reached significance, $F_{(1,77)} = 5.805$, $p = 0.0184$, because in the Nasalization by Tongue Height experiment the divided attention error rate differs more from the selective attention error rate than in the Pitch by Tongue Height experiment. As predicted by the hypothesis that the dimensions at one of the orientations are perceptually primary, the interaction of Rotation with Attention did approach significance, $F_{(2,77)} = 2.932$, $p = 0.0593$, reflecting progressive reduction of the difference in performance on the divided vs. selective attention tasks with rotation (fig. 3a).

However, figures 3b, c, which break out the effect of Rotation by Experiment, show that while rotation away from 0° uniformly increases the error rate in selective attention tasks in both experiments, its effect on error rates in the divided attention tasks is not uniform across or within experiments. Inexplicably, the divided attention error rate is highest at 22.5° for the Nasalization by Tongue Height experiment (fig. 3b), while that error rate drops with rotation to 22.5° and then with rotation to 45° rises again in the Pitch by Tongue Height experiment (fig. 3c). While the increase in error rate for selective attention with rotation away from 0° does support interpreting the dimensions aligned with the array at 0° as the perceptually primary ones, the changes in divided attention error rates do not. The assessment of perceptual primacy is therefore pursued along a different path in the evaluation of congruence in the next section.

3.4. Assessing Perceptual Primacy through Congruence

In both experiments, there is a contrast between stimuli in which dimensions vary congruently vs. those where they vary incongruently, but the two experiments differ in whether the congruent stimuli are also those which exhibit the naturally occurring covariation of dimensions. For the Nasalization by Tongue Height experiment, the congruent stimuli, B and C, are also those which exhibit the natural covariation of dimensions, a large $N_2{-}N_1$ difference with high F_1 and vice versa. For the Pitch by Tongue Height experiment, on the other hand, the congruent stimuli exhibit the opposite pattern of covariation, high F_0 with high F_1 and vice versa. Of interest here, then, is whether the congruent or the naturally covarying vowels are perceived more accurately in both experiments.

For each experiment, subject, rotation, and task type (baseline, correlated, selective attention, and divided attention), the errors were separately tabulated for each of A, B, C, and D, and then expressed as a proportion of the total responses for that subject, experiment, rotation and task type. The error proportions for the congruent stimuli, B and C, were then added together, as were the error proportions for the incongruent stimuli, A and D, and a congruence score was calculated by subtracting the congruent sum (B+C) from the incongruent one (A+D). These scores would tend to be positive if fewer errors were made for the putatively congruent B and C than the putatively incongruent A and D. Mean congruence scores (with standard errors) are presented in table 6 (note that errors are pooled within a task type).

Taking these differences for each Subject as the dependent variable, a three-way ANOVA was performed in which the independent variables were Experiment, Rotation, and Task Type.

Experiment was significant ($F_{1,161} = 22.708$, $p < 0.0001$) because congruence scores were predominantly negative in the Nasalization by Tongue Height experiment,

---

Table 6. Mean congruence score
errors for stimuli B and C and sumn
Nasalization by Tongue Height and

Task	Rotation
	0°
	mean
Nasalization by Tongue Height	
Baseline	–0.012
Correlated	–0.041
Selective	–0.097
Divided	–0.016
mean	–0.041
Pitch by Tongue Height	
Baseline	–0.010
Correlated	0.004
Selective	0.023
Divided	–0.025
mean	–0.002

but near zero or weakly positi
by Tongue Height experiment
and C were clearly perceived l
for the Nasalization by Tong
periment than incongruent
there was little difference in
tween congruent and incongru
the Pitch by Tongue Heigh
Task Type was also significan
negative congruence scores
and selective attention task
scores for divided attention a
no polarity for baseline tasks.
not significant as a main effe

Two interactions achievec
Experiment by Task Type, F
$p = 0.0132$ (fig. 4a), and Rot
Type, $F_{(6,161)} = 3.739$, $p = 0.$

' dimensions. For the Nasal-
ngue Height experiment, the
nuli, B and C, are also those
the natural covariation of di-
arge $N_2–N_1$ difference with
vice versa. For the Pitch by
ıt experiment, on the other
gruent stimuli exhibit the op-
of covariation, high F_0 with
vice versa. Of interest here,
er the congruent or the natu-
g vowels are perceived more
oth experiments.
xperiment, subject, rotation,
(baseline, correlated, selective
divided attention), the errors
y tabulated for each of A, B,
l then expressed as a propor-
al responses for that subject,
ıtation and task type. The er-
s for the congruent stimuli, B
hen added together, as were
ortions for the incongruent
l D, and a congruence score
by subtracting the congruent
from the incongruent one
scores would tend to be posi-
rrors were made for the puta-
nt B and C than the puta-
ent A and D. Mean congru-
vith standard errors) are pre-
ble 6 (note that errors are
a task type).
se differences for each Sub-
endent variable, a three-way
performed in which the inde-
bles were Experiment, Rota-
Type.
was significant ($F_{(1,161)}$=
.0001) because congruence
redominantly negative in the
y Tongue Height experiment,

Table 6. Mean congruence scores (with standard errors), as differences between summed proportion of errors for stimuli B and C and summed proportion of errors for A and D, i.e. [errors(A+D)−errors(B+C)]: Nasalization by Tongue Height and Pitch by Tongue Height

Task	Rotation							
	0°		22.5°		45°		mean	
	mean	SE	mean	SE	mean	SE	mean	SE
Nasalization by Tongue Height								
Baseline	−0.012	0.007	0.029	0.011	−0.019	0.014	0.000	0.008
Correlated	−0.041	0.020	−0.098	0.034	−0.051	0.036	−0.063	0.018
Selective	−0.097	0.019	−0.084	0.037	−0.093	0.028	−0.091	0.016
Divided	−0.016	0.034	−0.006	0.028	0.055	0.024	0.011	0.017
mean	−0.041	0.012	−0.040	0.017	−0.027	0.016	−0.036	0.009
Pitch by Tongue Height								
Baseline	−0.010	0.008	0.024	0.014	−0.001	0.007	0.004	0.006
Correlated	0.004	0.007	−0.010	0.010	0.004	0.010	−0.001	0.005
Selective	0.023	0.031	−0.056	0.024	0.009	0.036	−0.008	0.018
Divided	−0.025	0.027	0.040	0.018	0.096	0.029	0.037	0.017
mean	−0.002	0.011	−0.001	0.011	0.027	0.013	0.008	0.007

but near zero or weakly positive in the Pitch by Tongue Height experiment. Congruent B and C were clearly perceived less accurately for the Nasalization by Tongue Height experiment than incongruent A and D, but there was little difference in accuracy between congruent and incongruent stimuli in the Pitch by Tongue Height experiment. Task Type was also significant as a result of negative congruence scores for correlated and selective attention tasks vs. positive scores for divided attention and scores with no polarity for baseline tasks. Rotation was not significant as a main effect.

Two interactions achieved significance, Experiment by Task Type, $F_{(3,161)} = 3.690$, $p = 0.0132$ (fig. 4a), and Rotation by Task Type, $F_{(6,161)} = 3.739$, $p = 0.0016$ (fig. 4b). The Experiment by Task Type interaction leads to reinterpretation of the two significant main effects, since congruence scores for correlated and selective attention tasks are only markedly negative in the Nasalization by Tongue Height experiment; in the Pitch by Tongue Height experiment, the congruence scores in these two tasks are much closer to zero, as are congruence scores overall for this experiment. This disparity is reversed for divided attention where congruence scores are more strongly positive in the Pitch than Nasalization by Tongue Height experiment. In correlated and selective attention tasks, the two members of either the congruent and incongruent pairs must always be assigned to different classes. In the Nasalization by Tongue

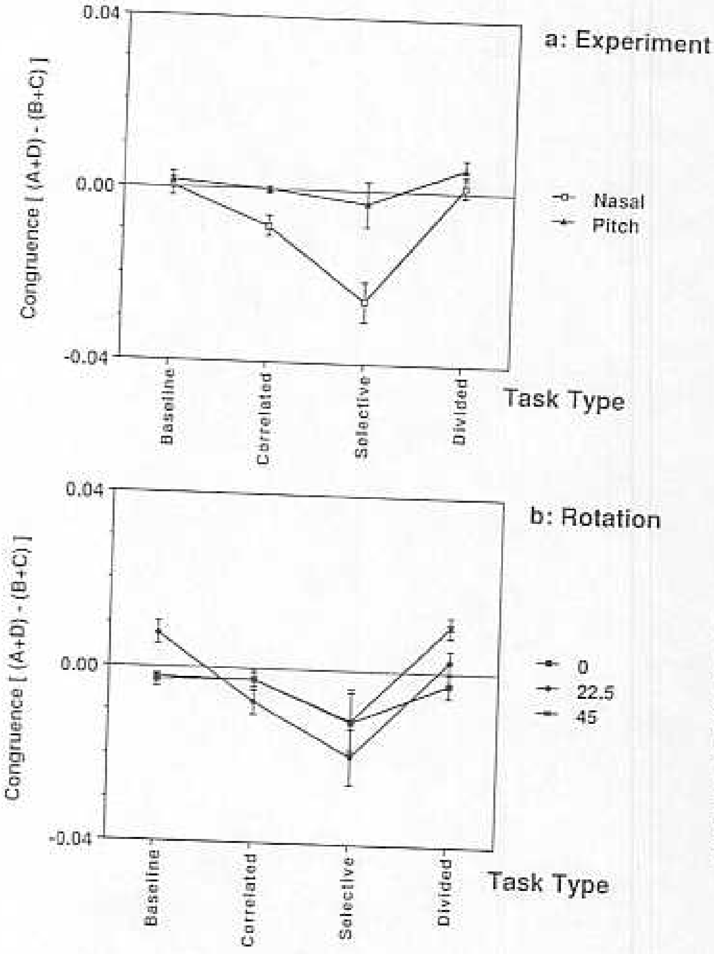

Fig. 4. Mean congruence scores (with standard errors) expressed as difference between the sum of the congruent from the sum of the congruent stimuli, [(A+D)−(B+C)]: the effects of Task Type on the Nasalization vs. Pitch by Tongue Height experiments (**a**) and the effects of Task Type on Rotation (**b**).

Height experiment, achieving this result is clearly much more difficult for B and C, the stimuli in which the two dimensions covary as in natural vowels, than A and D, in which they covary inversely, but only a much smaller disparity can be observed in the Pitch by Tongue Height experiment, and that is due to negative congruence scores at 22.5° alone. In the divided attention task, where the congruent stimuli are to be distinguished from the incongruent, there is an advantage for the congruent over the incongruent stimuli. This advantage is greater in the Pitch by Tongue Height experiment, where it increases with rotation away from 0°, than in the Nasalization by Tongue Height experiment, where congruence scores are only noticeably positive at 45°. Interpretation of both results is aided by figure 1. If the A-B and C-D sides of the stimulus array were lengthened perceptually with rotation away from 0°, performance should improve for the incongruent A and D stimuli over the congruent B and

---

C stimuli for the two tasks
members of each pair are assi
ent classes, the correlated an
tention tasks, but this advant
reversed for the task where
stimuli are assigned to the s
divided attention task. What
zling is why the first effect
nounced in the Nasalizatio
Height experiment, but the s
pronounced in the Pitch by
experiment.

A series of independent tw
were performed for each c
Task Type and Rotation in
ment to see whether the con
differed from zero; their re
those obtained from the A
Nasalization by Tongue Hei
significantly negative results
in 3 (of 12) tests ($p<0.05$),
cantly negative results
($0.05<p<0.10$), negative tr
ers, and positive values for o
line task at 22.5° ($t_{(7)}$ = 2.6
and the divided attention ta
2.246, $p = 0.0596$). The pr
negative congruence scores
putatively incongruent pai
and D, were usually classi
rately than the putatively co
and C, in this experiment. H
sult does not support the
Nasalization and Tongue
ceptually primary dimensio
not depend on the listeners'
independent measurements
along the two dimensions.
in which the dimensions co
are apparently more readil
This result is a generalizati
error rates observed for A a

Fig. 4. Mean congruence scores (with standard errors) expressed as difference between the sum of the congruent from the sum of the congruent stimuli, [(A + D) − (B + C)]: the effects of Task Type on the Nasalization vs. Pitch by Tongue Height experiments (**a**) and the effects of Task Type on Rotation (**b**).

ent stimuli. This advantage is
: Pitch by Tongue Height ex-
ere it increases with rotation
°, than in the Nasalization by
ht experiment, where congru-
re only noticeably positive at
tation of both results is aided
the A-B and C-D sides of the
y were lengthened perceptu-
ation away from 0°, perfor-
l improve for the incongruent
uli over the congruent B and

C stimuli for the two tasks in which the members of each pair are assigned to different classes, the correlated and selective attention tasks, but this advantage should be reversed for the task where the congruent stimuli are assigned to the same class, the divided attention task. What remains puzzling is why the first effect is more pronounced in the Nasalization by Tongue Height experiment, but the second is more pronounced in the Pitch by Tongue Height experiment.

A series of independent two-tailed t tests were performed for each combination of Task Type and Rotation in each Experiment to see whether the congruence scores differed from zero; their results resemble those obtained from the ANOVA. For the Nasalization by Tongue Height experiment, significantly negative results were obtained in 3 (of 12) tests ($p < 0.05$), almost significantly negative results in 2 others ($0.05 < p < 0.10$), negative trends in 5 others, and positive values for only 2: the baseline task at 22.5° ($t_{(7)} = 2.634$, $p = 0.0337$) and the divided attention task at 45° ($t_{(7)} = 2.246$, $p = 0.0596$). The preponderance of negative congruence scores shows that the putatively incongruent pair of stimuli, A and D, were usually classified more accurately than the putatively congruent pair, B and C, in this experiment. However, this result does not support the conclusion that Nasalization and Tongue Height are perceptually primary dimensions, since it does not depend on the listeners' ability to make independent measurements of the stimuli along the two dimensions. Instead, vowels in which the dimensions covary unnaturally are apparently more readily distinguished. This result is a generalization of the lower error rates observed for A and D in the negatively correlated task compared to B and C in the positively correlated one (see end of section 3.2). The t tests for the Pitch by Tongue Height experiment yielded quite different results. Congruence scores were significantly positive in just one case, the divided attention task at 45° ($t_{(7)} = 3.300$, $p = 0.0131$) and approached positive significance in one other, divided attention at 22.5° ($t_{(7)} = 2.218$, $p = 0.062$); there was 1 case where the test approached negative significance, selective attention at 22.5° ($t_{(7)} = -2.350$, $p = 0.051$; otherwise, the congruence scores did not differ from zero and no trends emerged (5 were weakly positive and 4 weakly negative). The overall failure of the congruence scores to differ significantly from zero in the Pitch by Tongue Height experiment suggests that listeners do no reliably classify vowels along these two dimensions independently of perceived vowel height – this result also generalizes the similarity in error rates between the positively and negatively correlated tasks in the Pitch by Tongue Height experiment.

3.5. Summary

That correlated variation facilitated classification, while orthogonal variation interfered with it in both these experiments suggests that both Nasalization and Pitch are integrated with Tongue Height in the perception of vowels. The perceptual primacy of any of these dimensions remains in doubt, however, since although selective attention performance did decline with rotation away from 0°, divided attention performance did not improve uniformly. Nasalization appears from these results to be more of an independent attribute of vowels than Pitch does, although the analysis of the congruence scores suggests a different interpretation.

There it was found that stimuli which combine Nasalization with Tongue Height in ways not observed in natural speech were classified more accurately than those that did, across tasks and rotations. Those stimuli in which these dimensions covary naturally are classified less accurately as a result of more thorough integration of the two dimensions in them, which makes it more difficult to independently evaluate their values along each dimension. It is harder to detect whether a vowel has a high amount of Nasalization when its Tongue Height is Low (= high F_1) and vice versa, than when these two dimensions covary in the opposite direction. With covariation in opposite directions (low F_1 with high N_2–N_1 and vice versa), two relatively different dimensions of difference can be more readily detected, and as a result those vowels with unnatural covariation are classified more accurately. The perceptual primacy of Nasalization and Tongue Height can thus be blocked by their integration into a percept of vowel height. On the other hand, by either test, Pitch does not appear to be a perceptually primary attribute of vowels in these stimuli.

4. Discussion

4.1. Implications for Theories of Speech Perception

These experiments address the question of whether the effects of independent articulations are separated in perception as predicted by direct realist theories, at least in their strong form. Proponents of these theories have argued that such separation is possible so long as a coarticulatory source is available [for nasalization and tongue height see Beddor et al., 1986, and Krakow et al., 1988; for similar results regarding pitch and tongue height see Silverman, 1987, and see Fowler and Smith, 1986], but they would argue that when no coarticulatory source exists, as with these stimuli, then integration may take place. But this does not explain why these patterns of covariation occur in natural speech outside contexts in which they could arise via coarticulation: low vowels have intrinsically more nasalization and lower pitches than high vowels. In fact, if it is the height of the tongue that is the target of perception, then these patterns of covariation are just so much noise, since in a direct realist theory, they convey the occurrence of irrelevant articulations.

On the other hand, from the perspective that all the speaker's articulatory behavior is intended to influence the perceived height of F_1, and from that to produce a particular percept of vowel height, then the observed covariation AND the ready integration across dimensions become intelligible. The objects of speech perception are not articulatory, but acoustic, and the many-to-one mapping of the one onto the other ceases to be a problem and becomes a positive advantage for the listener. The integration between Nasalization or Pitch and Tongue Height demonstrated in these experiments supports the strong version of the auditory enhancement theory [Diehl and Kluender, 1989; Diehl, this volume], although these experiments do not point to precisely what auditory mechanism is responsible [for a likely mechanism, see Syrdal and Gopal, 1986].

4.2. Implications for the Representation of Vowel Height

4.2.1. Mutual Dependencies between the Covarying Articulations and Their Categorical Representation

The results of the two experiments also undermine the traditional phonological representation of vowel height as a single articulatory property. At least from the listener's

---

point of view, vowel height c
integrated rather than additiv
the effects of tongue height,
and rate of vocal fold vibrati
ers may integrate the effects c
varying supralaryngeal arti
The speaker does not achiev
covariation because raising t
particular height in the mout
perturbs the other articulato
trolling each of these articula
dependently. The linked con
ticulation, the observed cova
have a single goal, but it is ac
properly perceptual, the m
perceived F_1 and through it
height of the vowel.

This covariation raises
problem for the representa
height in phonological contra
of these covarying articulato
independently in the statem
logical contrasts or rules, the
els differing in height also b
the features [high] and [low] (
[nasal], [High] (pitch), as w
[advanced tongue root], [rou
the synchronic and diachro
vowels differing in height su
values for these other articul
ways so constrained as to re
phonemic. For example:

(1) *ATR*. Differences in t
sition have brought about ch
height in Mon-Khmer lang
son, 1976; Huffman, 1976].
languages which exhibit vow
tongue root position in W
Africa, the low vowel does n
instead takes a single value,
[Hall et al., 1974]. Some
quently, the high vowels al

ich they could arise via coar-
w vowels have intrinsically
tion and lower pitches than
n fact, if it is the height of the
the target of perception, then
of covariation are just so
ince in a direct realist theory,
ie occurrence of irrelevant ar-

er hand, from the perspective
eaker's articulatory behavior
to influence the perceived
and from that to produce a
cept of vowel height, then the
riation AND the ready inte-
dimensions become intelligi-
ts of speech perception are
ry, but acoustic, and the
napping of the one onto the
be a problem and becomes a
tage for the listener. The inte-
n Nasalization or Pitch and
t demonstrated in these ex-
orts the strong version of the
ncement theory [Diehl and
9; Diehl, this volume], al-
xperiments do not point to
auditory mechanism is re-
a likely mechanism, see Syr-
, 1986].

ions for the Representation
ight

l Dependencies between
g Articulations and
rical Representation

of the two experiments also
traditional phonological rep-
owel height as a single artic-
y. At least from the listener's

---

point of view, vowel height contrasts are an integrated rather than additive composite of the effects of tongue height, velum height, and rate of vocal fold vibration, and listeners may integrate the effects of the other covarying supralaryngeal articulators, too. The speaker does not achieve the observed covariation because raising the tongue to a particular height in the mouth mechanically perturbs the other articulators, but by controlling each of these articulatory events independently. The linked control of each articulation, the observed covariation, DOES have a single goal, but it is acoustic or more properly perceptual, the manipulation of perceived F_1 and through it the perceived height of the vowel.

This covariation raises an interesting problem for the representation of vowel height in phonological contrasts. Since each of these covarying articulators is referred to independently in the statement of phonological contrasts or rules, then should vowels differing in height also be specified for the features [high] and [low] (tongue height), [nasal], [High] (pitch), as well perhaps as [advanced tongue root], [round], etc.? Both the synchronic and diachronic behavior of vowels differing in height suggest that their values for these other articulators are not always so constrained as to remain safely subphonemic. For example:

(1) *ATR.* Differences in tongue root position have brought about changes in vowel height in Mon-Khmer languages [Gregerson, 1976; Huffman, 1976]. In many of the languages which exhibit vowel harmony for tongue root position in West and East Africa, the low vowel does not contrast, but instead takes a single value, usually [−ATR] [Hall et al., 1974]. Somewhat less frequently, the high vowels also do not contrast, typically taking [+ATR] values [Hall et al., 1974].

(2) *Rounding.* Contrasts for lip rounding are similarly dependent on tongue height, where in a great many languages, only nonlow back vowels are rounded. In languages with front rounded vowels, the occurrences of such a vowel at a given height implies the occurrence of all front rounded vowels of greater height, though not vice versa [Crothers, 1978; Disner, 1983, 1984, 1986]. Hayes [1990] describes diphthongization in Quebec French, the Lund dialect of Swedish, and Eastern Finnish in which a monophthong breacks into higher and lower components. In each of these cases, if a nonlow back vowel becomes low, it loses distinctive rounding, and if a low back vowel becomes nonlow it acquires it. The case of Quebec French is particularly interesting since a contrast for backness is preserved between the low vowels derived by diphthongization despite the loss of rounding from the back one, which shows that rounding is linked to height in this process.

(3) *Nasal.* The susceptibility of vowels to nasalization depends on their height: low vowels are more likely to be distinctively nasalized than nonlow ones, and high vowels more than mid ones [Ohala, 1974, 1975; Ruhlen, 1978]; nasalization also frequently changes vowel height, both diachronically and synchronically, lowering high and mid vowels and raising low ones [Ruhlen, 1978; Wright, 1980; Beddor, 1983; Beddor et al., 1986].[1]

---

[1] Neither favoring nasalized high vowels over nasalized mid vowels nor raising nasalized low vowels is predicted by the uniform (and inverse) covariation of nasalization with vowel height. However, the acoustic effects, and their likely perceptual con-

(4) *Pitch.* Differences in rate of fundamental frequency between high and low vowels have led in one case to tone splitting, in the Northern Mon-Khmer language U [Svantesson, 1989], although this is very rare [Hombert, 1977a],[2] while in Foochow and Lahu, vowels have been raised by high and rising tones, respectively [Hombert, 1977a]. Andersen [1986] suggests that [+ATR] vowels may have raised High tones historically in the Chari-Nile language, Lugbara, so perhaps vowel height may influence the phonological use of F_0 through another of its covarying articulations.

Many of these phonological consequences are the result of what Ohala [1981] has described as the misperception of the speaker's intent by the listener. For example, nasalization changes vowel height because its acoustic effects lead the listener to hear a vowel of different height than the speaker actually produced, and the listener's (or the language learner's) own production is determined henceforth by that misperception. But it is also important to distinguish this sort of case from others listed above, where it is just the inherent covariation among articulations, and their convergent perceptual effects, rather than the effects of one distinctive feature on the perception of another, that influence phonological patterning. Inherent covariation among articulations acts more as a constraint on phonological inventories and processes than a source of change. Even so, the splitting of U's tones as a result of inherent difference in F_0 between high and low vowels shows that change, though rare, can come from this inherent covariation.

Representing these articulations, or their perceptual correlates, in terms of individual features does not require that they be separable, however, since the value for one feature is still predictable from the value for another, and this covariation could be achieved by means of bidirectional redundancy rules of the form $[\alpha \text{ high}] \longleftrightarrow [-\alpha \text{ nasal}]$, etc. The bidirectionality of these rules expresses the perceptual integration of the acoustic consequences of these articulations, whose independence, strictly from the speaker's point of view, is expressed by the use of a separate feature for each one. (Stevens et al. [1986] and Stevens and Keyer [1989] give similar arguments for using distinctive features for enhancement purposes,

---

sequences, of the light nasalization that accompanies low oral vowels are different from the heavy nasalization of nasalized ones. Lowering the velum both replaces the first oral pole with a nasal pole–nasal zero–oral pole triplet and raises the frequency of the oral pole. Raising the frequency of the oral pole would appear to be the predominant effect of light nasalization, since the nasal zero is still close enough to the nasal pole to largely cancel it out, and would make an already low vowel sound lower, as well as making a higher vowel sound lower. Increasing nasalization augments the separation of the nasal zero from the nasal pole, and may also reduce the spectral prominence of the oral pole. With low vowels, this has the effect of shifting the center of gravity in that part of the spectrum to a lower frequency, like that of a higher vowel. For high vowels, on the other hand, these effects are still likely to make the vowel sound lower, especially in the highest ones in which the nasal pole is actually above the oral pole.

[2] Hombert [1977a], who argues that F_0 differences between high and low vowels rarely affect tones, cites just four cases, Hausa, Middle Chinese, the Omei dialect of Mandarin, and Ngizim. According to Hombert, the Hausa and Middle Chinese cases can and probably should be analyzed differently, and in Ngizim, a low vowel [a] in the first syllable of a verb leads unexpectedly to the verb having a high tone rather than low tone. This leaves just the single case of the Omei dialect of Mandarin beside the Mon-Khmer language U.

---

but Stevens and Keyser disting
sonants, between what they
features, those employed to b
and 'secondary' features, w
these contrasts. Here, I argu
distinction, at least for vowels

Two further pieces of evic
categorical representation of
ing articulations. First are c
splitting in U [Svantesson, 19
inherent difference in pitch
and low vowels has split earl
higher and lower reflexes ir
day language. The pitch d
clearly become categorical,
so by itself, even if it were n
originally. (The neutralizati
contrasts in high and low vo
languages [Hall et al., 1974
dence of the categorical beh
covarying articulations.) Se
dence that some of these ar
not covary in step with tong
instead show extreme 'high'
vowels above some arbitrary
in the tongue height contin
treme 'low' values below it.
categoricality is apparent in
data on lip rounding and is
istic of pitch differences as
tongue height [see Silverman,
cussion]. The partitioning
height continuum by the extr
the covarying articulations
their use by the speaker to e
trasts, or more generally t
motivations for their use, bu
viewed, from the perspective
mar of the language, as the
categoricality from phonolo
tations [see Pierrehumbert, 1
evidence that phonetic i

these phonological conse-
he result of what Ohala [1981]
d as the misperception of the
ent by the listener. For exam-
ion changes vowel height be-
istic effects lead the listener to
l of different height than the
ially produced, and the lis-
e language learner's) own pro-
etermined henceforth by that
n. But it is also important to
his sort of case from others
where it is just the inherent co-
iong articulations, and their
erceptual effects, rather than
one distinctive feature on the
f another, that influence pho-
terning. Inherent covariation
ilations acts more as a con-
nological inventories and pro-
source of change. Even so, the
's tones as a result of inherent
F_0 between high and low vow-
at change, though rare, can
is inherent covariation.
ng these articulations, or their
rrelates, in terms of individual
not require that they be sepa-
r, since the value for one fea-
redictable from the value for
this covariation could be
neans of bidirectional redun-
f the form [α high] ⟷ [−α na-
bidirectionality of these rules
perceptual integration of the
sequences of these articula-
independence, strictly from
point of view, is expressed by
separate feature for each one.
[1986] and Stevens and Keyer
nilar arguments for using dis-
es for enhancement purposes,

---

but Stevens and Keyser distinguish, for consonants, between what they call 'primary' features, those employed to bear a contrast, and 'secondary' features, which enhance these contrasts. Here, I argue against this distinction, at least for vowels.)

Two further pieces of evidence support categorical representation of these covarying articulations. First are cases like tone splitting in U [Svantesson, 1989], where the inherent difference in pitch between high and low vowels has split earlier tones into higher and lower reflexes in the present-day language. The pitch difference has clearly become categorical, and has done so by itself, even if it were not categorical originally. (The neutralization of [ATR] contrasts in high and low vowels in many languages [Hall et al., 1974] is also evidence of the categorical behavior of these covarying articulations.) Second is evidence that some of these articulations do not covary in step with tongue height, but instead show extreme 'high' values for all vowels above some arbitrary dividing point in the tongue height continuum and extreme 'low' values below it. This incipient categoricality is apparent in Linker's [1982] data on lip rounding and is also characteristic of pitch differences associated with tongue height [see Silverman, 1987, for discussion]. The partitioning of the tongue height continuum by the extreme values of the covarying articulations originates in their use by the speaker to exaggerate contrasts, or more generally the perceptual motivations for their use, but can also be viewed, from the perspective of the grammar of the language, as the persistence of categoricality from phonological representations [see Pierrehumbert, 1980, for other evidence that phonetic implementation needs to refer directly to the categorical properties of speech sounds].

4.2.2. Does Covariation Persist When an Articulation Is Used Contrastively?

However, representing all these covarying articulations in terms of bidirectional redundancy rules referring to categorical features also overstates the magnitude of the differences in these articulations between high and low vowels, and even worse incorrectly predicts that these features may not contrast phonemically in vowels. For example, if low vowels are inherently [+ nasal], then no contrast should be possible between a low oral and low nasal vowel; for that matter, if each height has an associated specification for nasality, then no vowel, at any height, should contrast for nasality. What is needed here is a phonetic representation that allows specification of nasality differences among what are phonemically all [− nasal] vowels.

The solution to this dilemma is to recognize that vowel height contrasts are abstract in the phonology;[3] the point of this paper has been in fact to argue that they are so abstract in that component of the grammar as not to have any necessary association with the height of the tongue *per se* (nor with degree of constriction). (Whether other phonemic contrasts are abstract in the same sense remains to be determined, but I su-

---

[3] John McCarthy [personal commun.] notes that this view is very like the traditional view of stress in English, as an abstract coding of prominence, which is realized in terms of a variety of different articulations, variously selected according to context [Beckman, 1986]. Stress, too, may exhibit integration of different dimensions, for example of amplitude and duration, as suggested by Beckman [1986] in her notion of 'total amplitude'.

spect they are.) Once these abstract height contrasts are implemented in the phonetics, however, they are spelled out in terms of a set of at least quasi-categorical articulations involving the velum, larynx, jaw, lips, and tongue. If another phonemic contrast also gets spelled out in terms of any of these articulators, then in some cases their inherent covariation in spelling out vowel height will be subordinate to, and in physical terms smaller than the articulatory events and acoustic consequences that spell out the phonemic contrast. That superordinate contrast may even wipe out the inherent covariation, as for example if phonemically nasalized vowels of different heights (or those which are nasalized by an adjacent nasal) no longer differ in nasalization, as phonemically oral vowels do.

The data show, however, that differences in nasalization between vowels of different heights are frequently preserved when they're contextually nasalized, and in some cases even when they are distinctively nasalized. Clumeck [1976] shows that relative to the duration of the vowel, the higher the vowel the later the velum begins to lower in oral vowels before nasal consonants in English,[4] French, Amoy, and Hindi, and he finds the same dependency in Swedish long but not short vowels, which are so short that the whole vowel is nasalized. The magnitude of velar lowering in oral vowels before nasal consonants was not entirely predictable from height for all the speakers, but even so all the significant differences between vowels rank them in terms of magnitude of velar lowering as: low > mid > high. Moll [1962] and Kuehn [1976] also report larger velar opening for low than high vowels before nasals. In addition, Clumeck [1976] observed that low vowels were produced by English speakers with velar lowering in oral environments, but this was not observed for the other languages [this difference in velum height between high and low vowels outside of nasal contexts was also observed by Moll, 1962, and Kuehn, 1976, for English, and by Kuenzel, 1977, for isolated German vowels]. Finally, Clumeck [1976] reports that the magnitude of velar lowering did not differ for distinctively nasalized vowels of different heights in Hindi or French, but for Amoy and Portuguese it did, with less velar lowering for higher vowels in both languages. The data of Bell-Berti et al. [1979] from a single speaker of English show consistently lower velum position in low than high vowels in oral as well as nasal contexts in disyllabic utterances of the form [fVCmVp] and [fVmCVp], where the C is an oral obstruent and the Vs were either both [i] or both [a]. The velum was actually lower in all the segments, consonants as well as vowels, oral as well as nasal consonants, when the vowel was [a]. Differences in nasalization associated with vowel height must not only survive contextual nasalization but also influence its contrastive realization. Finally, contrary to all these observations, Vaissiere [1988] found no differences in degree of velar lowering between

---

[4] Cohn [1990] argues that there is no phonological rule which uniformly nasalizes vowels before nasals in English. If the vowel does not actually come to bear a [+ nasal] specification, vowels of different heights could still differ in degree of nasalization in that context. However, even Cohn [1990] notes that deletion or severe truncation of the nasal consonant's oral articulation renders a vowel [+ nasal] in a word such as *can't*, so it should still be possible to test for the persistence of height-related differences in nasalization in the presence of positive specification for that feature. I am not aware of any study that presents the relevant data, unfortunately.

---

vowels of different heights f
English in connected speech
Even an independent ph
on a vowel for an articulati
wise covaries with tongue
always constrain its use in c
height contrasts, and in the
ally exaggerates its effect. L
man [1984] have shown tha
vowels contrasting in vowel
cus, the difference in F_0 be
low vowels is far larger tha
not [similar results are pre
lish by Steele, 1986]. Focus
implemented in part by map
cent containing a High tone
in question, so F_0 is being us
in this case (and also in
data). Another possible as
local expansion of the pitch
contribute substantially to
the F_0 differences between
vowels in this context [Beck
humbert, 1986, and Pier
Beckman, 1988]. Convers
[1987] has shown that list
the expected effects of vo
when judging whether a vo
relatively high F_0 on a high
uted to the vowel's height
cient for the vowel to be j
while a relatively low F_0 o
sufficient for it to be ju
Clearly, F_0 can at once co
veying both vowel height a
F_0 can probably be us
poses at once because the
an F_0 value in terms of the
sent intonational contrasts i
depends on F_0 values else
contour, and perhaps eve
contour. Its local value in

ort vowels, which are so short
ole vowel is nasalized. The mag-
elar lowering in oral vowels be-
consonants was not entirely pre-
om height for all the speakers,
o all the significant differences
wels rank them in terms of mag-
elar lowering as: low > mid >
[1962] and Kuehn [1976] also re-
velar opening for low than high
ore nasals. In addition, Clumeck
rved that low vowels were pro-
nglish speakers with velar lower-
environments, but this was not
or the other languages [this dif-
velum height between high and
outside of nasal contexts was
ed by Moll, 1962, and Kuehn,
nglish, and by Kuenzel, 1977, for
rman vowels]. Finally, Clumeck
rts that the magnitude of velar
d not differ for distinctively na-
els of different heights in Hindi
out for Amoy and Portuguese it
s velar lowering for higher vow-
anguages. The data of Bell-Berti
from a single speaker of English
stently lower velum position in
gh vowels in oral as well as na-
in disyllabic utterances of the
Vp] and [fVmCVp], where the C
bstruent and the Vs were either
oth [a]. The velum was actually
l the segments, consonants as
els, oral as well as nasal conso-
the vowel was [a]. Differences
on associated with vowel height
ly survive contextual nasaliza-
o influence its contrastive real-
lly, contrary to all these obser-
ssiere [1988] found no differ-
gree of velar lowering between

vowels of different heights for 2 speakers of English in connected speech.

Even an independent phonemic contrast on a vowel for an articulation which otherwise covaries with tongue height does not always constrain its use in conveying vowel height contrasts, and in the case of F_0 actually exaggerates its effect. Ladd and Silverman [1984] have shown that when German vowels contrasting in vowel height are in focus, the difference in F_0 between high and low vowels is far larger than when they are not [similar results are presented for English by Steele, 1986]. Focus was apparently implemented in part by mapping a pitch accent containing a High tone onto the vowels in question, so F_0 is being used contrastively in this case (and also in Steele's English data). Another possible aspect of focus, a local expansion of the pitch range, may also contribute substantially to the increase in the F_0 differences between high and low vowels in this context [Beckman and Pierrehumbert, 1986, and Pierrehumbert and Beckman, 1988]. Conversely, Silverman [1987] has shown that listeners factor out the expected effects of vowel height on F_0 when judging whether a vowel is in focus: a relatively high F_0 on a high vowel is attributed to the vowel's height and is insufficient for the vowel to be judged in focus, while a relatively low F_0 on a low vowel is sufficient for it to be judged in focus. Clearly, F_0 can at once contribute to conveying both vowel height and intonation.

F_0 can probably be used for two purposes at once because the interpretation of an F_0 value in terms of the tones that represent intonational contrasts in the phonology depends on F_0 values elsewhere in the F_0 contour, and perhaps even on the entire contour. Its local value in isolation of this larger context is only relevant to judgments of vowel height (and perhaps obstruent voicing). Nasalization, on the other hand, may be used for only one purpose at a time because its contrastive use and its use in connection with vowel height are confined to the same local domain in most languages. This account predicts that F_0 would be constrained in its use in conveying vowel height contrasts in tone languages where F_0 contrasts in each vowel. This prediction is falsified by the partial maintenance of height differences in F_0 in the tone languages Yoruba [Hombert, 1977b] and Taiwanese [Zee, 1980]. In both languages, F_0 is substantially higher for high than low vowels produced on a high tone, but on a low tone the F_0 differences are negligible. This partial preservation probably reflects differences in the greater expandability of the high as compared to low end of the pitch range which have been observed elsewhere [Pierrehumbert and Beckman, 1988], and suggests that Ladd and Silverman's [1984] results should be reinterpreted as an effect of High tone rather than focus. In any case, even partial maintenance of height-related differences in F_0 when it is used contrastively is sufficient to show that contrastive use of an articulation on the individual vowel itself does not preclude its use in conveying vowel height.

However, despite the fact that the domain of phonological contrast for tone is as local in these tone languages as the domain of nasalization in most languages, the interpretation of F_0 values as particular tones probably still requires global reference to the F_0 contour, just as in languages where the domain of tonal contrasts is the intonational phrase. This hypothesis is supported by the fact that the phonetic implementation of tonal contrasts, however they arise in the phonology of different languages, is achieved by essentially the same mechanisms [Beckman and Pierrehumbert, 1986; Pierrehumbert and Beckman, 1988, chapter 8]. Interpretation of F_0 values in terms of tones, whether lexical or intonational, must therefore depend on the same global reference to the larger F_0 contour. That F_0 can be used in tonal as well as vowel height contrasts at the same time, while nasalization cannot, is because the PHONETIC domain of F_0 differences is far more global than that of nasalization; the phonological origin of the F_0 differences does not matter.

4.2.3. Phonetic Underspecification

The image of vowel articulations described above, in which all articulators contribute to conveying a contrast in vowel height through their influence on perceived F_1, differs markedly from the underspecified phonetic representations argued for by Keating [1988, 1990]. In conveying vowel height, no articulator lacks specification, although as just described, the realization of other contrasts by means of the same articulators may alter their contribution to conveying vowel height contrasts. A strong interpretation of Keating's [1988, 1990] proposal predicts coarticulation will arise only when an adjacent segment is unspecified for an articulation and allows that articulation in its neighbor to be anticipated or to carry over. This sort of coarticulation onto unspecified segments will only be possible if that articulations is not already used in the target segment to enhance its distinctiveness. However, the use of an articulation in conveying vowel height does not block an adjacent segment from imposing its value for that articulation on a vowel, as nasals impose [+ nasal] on preceding vowels, regardless of their height (anymore than the use of nasalization to enhance vowel height contrasts blocks nasal contrasts in vowels). The persistence of differences in nasalization between vowels of different heights under contextual if not contrastive nasalization suggests instead that coarticulation can take place even when the target sement is specified for that articulation, and that that specification alters the extent in both time and space of the coarticulation [see Stevens, 1990, for similar arguments].

4.2.4. The Phonological Equivalence of Covarying Articulations: [ATR] and [high]

The extremely rich representation of vowel height advocated here, which combines tongue height with both laryngeal and supralaryngeal articulations, is anticipated by Hyman's [1988] argument that [+ high, – low, – ATR] vowels must be phonetically equivalent to [– high, – low, + ATR] vowels if their behavior in vowel height harmony in Esimbi is to be accounted for. The results of Khabanyane's [1989] study of height assimilation in Sesotho [Southern Sotho, see also Doke and Mofokeng, 1957, and Clements, 1989] also become more intelligible from this perspective. In this language, which distinguishes apparently between five degrees of height (table 7), the mid vowels /ɛ/ and /ɔ/ are raised to /e/ and /o/ if the vowel in the following syllable is higher, i.e., by any of /e, o, ι, ʊ, o, u/, and /ι/ and /ʊ/ are raised, but not quite to /i/ and /u/, by either of /i, u/ (a following velar nasal /ŋ/, which derives historically from **ni*, also induces both kinds of raising). The change in /ɛ, ɔ/ to /e, o/ is apparently of [– ATR] to [+ ATR], but only some

---

of the conditioning vowels
[+ ATR], while the others
[+ high]. (The change in [ι, ʊ
tion of [+ ATR], but does
way.) The apparent lack o
the conditioning environm
once it is recognized that
with tongue height, in additi
a distinctive feature in this l
/ι, ʊ/ will have relatively hi
netically for [ATR] because
vowels and should trigger
[+ ATR] in /ɛ, ɔ/ just as th
[+ ATR] but nonhigh /e, o/
Esimbi, [+ high] appears to b
[+ ATR]; this is so becaus
narily covaries directly with
That the assimilation opera
[ATR] is also suggested by it
application to Sesotho vov
were what the rule manipula
should be raised toward /ι,
ing higher vowel [Clements,
that the raising of the mid
/ɛ, ɔ/ to /e, o/ collapses a
that of the high ones does
dental consequence of a rece
eralization of the occurrenc
contexts in which raising d
tion their occurrence, and v
therefore phonemically disti

However, stating the rul
from [– ATR] to [+ ATR] ov
fects on the high [– ATR] vo
not collapse with /i, u/. F
therefore be seen as a phone
vancing the tongue root of
they are followed by vowels
root is phonetically mo
whether because of phonolo
tion as [+ ATR] or because t
Advancing the tongue root c

se [+ nasal] on preceding vow-
ss of their height (anymore than
nasalization to enhance vowel
trasts blocks nasal contrasts in
e persistence of differences in
between vowels of different
ler contextual if not contrastive
suggests instead that coarticu-
take place even when the target
ecified for that articulation, and
ecification alters the extent in
nd space of the coarticulation
, 1990, for similar arguments].

e Phonological Equivalence
ying Articulations:
d [high]

remely rich representation of
t advocated here, which com-
e height with both laryngeal and
eal articulations, is anticipated
s [1988] argument that [+ high,
R] vowels must be phonetically
o [– high, – low, + ATR] vowels
avior in vowel height harmony
to be accounted for. The results
ane's [1989] study of height as-
n Sesotho [Southern Sotho, see
and Mofokeng, 1957, and Cle-
] also become more intelligible
perspective. In this language,
nguishes apparently between
of height (table 7), the mid
and /ɔ/ are raised to /e/ and
wel in the following syllable is
by any of /e, o, ι, ʊ, o, u/, and
are raised, but not quite to /i/
either of /i, u/ (a following
/ŋ/, which derives historically
so induces both kinds of rais-
nge in /ɛ, ɔ/ to /e, o/ is appar-
TR] to [+ ATR], but only some

---

of the conditioning vowels /e, o, i, u/ are [+ ATR], while the others /ι, ʊ, i, u/ are [+ high]. (The change in [ι, ʊ] is in the direction of [+ ATR], but does not go all the way.) The apparent lack of uniformity in the conditioning environment disappears once it is recognized that [ATR] covaries with tongue height, in addition to serving as a distinctive feature in this language. Thus, /ι, ʊ/ will have relatively high values phonetically for [ATR] because they are high vowels and should trigger the change to [+ ATR] in /ɛ, ɔ/ just as the contrastively [+ ATR] but nonhigh /e, o/ do. Just as in Esimbi, [+ high] appears to be equivalent to [+ ATR]; this is so because [ATR] ordinarily covaries directly with tongue height. That the assimilation operates in terms of [ATR] is also suggested by its discontinuous application to Sesotho vowels. If height were what the rule manipulated, then /e, o/ should be raised toward /ι, ʊ/ by a following higher vowel [Clements, 1989]. The fact that the raising of the mid [– ATR] vowel /ɛ, ɔ/ to /e, o/ collapses a contrast, while that of the high ones does not, is an accidental consequence of a recent, limited generalization of the occurrence of /e, o/ to contexts in which raising does not condition their occurrence, and where they are therefore phonemically distinct from /ɛ, ɔ/.

However, stating the rule as a change from [– ATR] to [+ ATR] overstates its effects on the high [– ATR] vowels, which do not collapse with /i, u/. Raising should therefore be seen as a phonetic process advancing the tongue root of vowels when they are followed by vowels whose tongue root is phonetically more advanced, whether because of phonological specification as [+ ATR] or because they are higher. Advancing the tongue root causes the nonhigh [– ATR] vowels to overlap completely their [+ ATR] counterparts, but only causes high [– ATR] vowels to approximate their [+ ATR] counterparts, perhaps because differences in tongue root position between [– ATR] and [+ ATR] vowels increase with tongue height. The iterative nature of this process, whereby a vowel which has undergone it can induce it in a [– ATR] vowel to its left, does not preclude it from being phonetic either, since other coarticulatory processes may apply in similar ways. For example, according to Ghazeli [1977], segments in the vicinity of a distinctively uvularized, 'emphatic' coronal in Tunisian Arabic are phonetically uvularized. If a nonuvularized coronal is encountered, it becomes fully uvularized and propagates uvularization even further than it would have gone otherwise (I am grateful to John McCarthy for making me aware of this example).

Table 7. Sesotho vowels, after Khabanyane [1989] and Clements [1989]

			High	Low	ATR
i		u	+	–	+
ι		ʊ	+	–	–
e		o	–	–	+
ɛ		ɔ	–	–	–
	a		–	+	–

4.3. Conclusion

The surface representation of vowel height, at the point of entry into phonetic implementation, would be in terms of the array of features whose settings will influence F_1, specifying the full panoply of articulations that covary in conveying this contrast. It is probable that no distinction can be drawn in this representation between distinctive and redundant features, all are on equal footing because they all contribute to conveying the same contrast. This rich specification is motivated perceptually, by the listener's demands for differences in F_1 enhanced by these covarying articulations. Finally, phonetic representations resemble phonological ones in using categorical specifications of articulatory events.

Acknowledgments

I am grateful to Carol Krumhansl and Paul Iverson for timely and useful discussions of method, to Lyn Frazier, John McCarthy, and Elisabeth Selkirk for their insights into the issues raised here and their requests for clarification, and to Randy Diehl and Keith Kluender for provoking these experiments in the first place.

References

Andersen, T.: Tone splitting and vowel quality: evidence from Lugbara. Stud. Afr. Ling. *17:* 55–68 (1986).

Beckman, M. E.: Stress and non-stress accent. Netherlands phonet. Archs vol. 7 (Foris, Dordrecht 1986).

Beckman, M. E.; Pierrehumbert, J. B.: Intonational structure in Japanese and English. Phonol. Yb. *3:* 255–310 (1986).

Beddor, P. S.: Phonological and phonetic effects of nasalization on vowel height (Indiana University Linguistics Club, Bloomington 1983).

Beddor, P. S.; Krakow, R. A.; Goldstein, L.: Perceptual constraints and phonological change: a study of nasal vowel height. Phonol. Yb. *3:* 197–217 (1986).

Bell-Berti, F.: An electromyographic study of velopharyngeal function in speech. J. Speech Hear. Res. *19:* 225–240 (1976).

Bell-Berti, F.; Baer, T.; Harris, K.; Niimi, S.: Coarticulatory effects of vowel quality on velar function. Phonetica *36:* 187–193 (1979).

Clements, G. N.: On the representation of vowel height (Cornell University, Ithaca, unpublished 1989).

Clumeck, H.: Patterns of soft palate movement in six languages. J. Phonet. *4:* 337–351 (1976).

Cohn, A. C.: Phonetic and phonological rules of nasalization. UCLA Working Papers Phonet. 76 (1990).

Crothers, J.: Typology and universals of vowel systems; in Greenberg, Universals of human language, V. II. Phonology, pp. 93–152 (Stanford University Press, Stanford 1978).

Diehl, R. L.; Kluender, K. R.: The objects of speech perception. Ecological Psychol. *1:* 121–144 (1989).

Diehl, R. L; Walsh, M. A.: An auditory basis for the stimulus-length effect in the perception of stops and glides. J. acoust. Soc. Am. *85:* 2154–2164 (1989).

Disner, S. F.: Vowel quality: the relation between universal and language-specific factors. UCLA Working Papers Phonet. *58* (1983).

Disner, S. F.: Insights on vowel spacing; in Maddieson, Patterns of sounds, chapter 9, pp. 136–155 (Cambridge University Press, Cambridge 1984).

Disner, S. F.: On describing vowel quality; in Ohala, Jaeger, Experimental phonology, pp. 69–79 (Academic Press, Orlando 1986).

Doke, C. M.; Mofokeng, S. M.: Textbook of Southern Sotho grammar (Longman, London 1957).

Fowler, C. A.: An event approach to the study of speech perception from a direct realist perspective. J. Phonet. *14:* 3–28 (1986).

Fowler, C. A.: Sound-producing sources as the objects of perception: rate normalization and nonspeech perception. J. acoust. Soc. Am. *88:* 1236–1249 (1990).

Fowler, C. A.; Smith, M. R.: Speech perception as 'vector analysis': an approach to the problems of invariance and segmentation; in Perkell, Klatt, Invariance and variability in speech processes, pp. 123–139 (Erlbaum, Hillsdale 1986).

Fox, R. A.; Trudeau, M. D.: A n scaling study of esophageal vo *45:* 30–42 (1988).

Gandour, J.; Weinberg, B.: On the tween vowel height and fundam evidence from esophageal speec 344–354 (1980).

Garner, W. R.: The processing of i structure (Erlbaum, Potomac 19

Ghazeli, S.: Back consonants and b lation in Arabic; PhD diss. Uni at Austin (1977).

Grau, J. W.; Kemler Nelson, D. G.: between integral and separable dence for the integrality of pitch exp. Psychol. general *117:* 347–

Gregerson, K.: Tongue-root and r Khmer; in Jenner, Thompson, S asiatic Studies, part 1. Oceanic cial Publication 13, pp. 323– Press of Hawaii, Honolulu 1976

Hall, B.; Hall, R.; Pam, M.; Myers Cherono, G.: African vowel ha from the vantage point of Kalenj see *57:* 241–267 (1974).

Hayes, B.: Diphthongization and c nology *7:* 31–71 (1990).

Hombert, J.-M.: Development of to height? J. Phonet. *5:* 9–16 (1977

Hombert, J.-M.: Consonant types, and tone in Yoruba. Stud. Afr. L (1977 b).

Honda, K.: Relationship between pi vowel articulation; in Bless, Ab physiology: contemporary resear issues, pp. 286–297 (College-H Diego 1987).

Honda, K.; Fujimura, O.: Phonolog cal explanation – intrinsic vo phrasal declination, 6th Vocal Conf., Stockholm 1989.

House, A. S.; Fairbanks, G.: The in sonant environment upon the se tic characteristics of vowels. J. ac *25:* 105–113 (1953).

House, A. S.; Stevens, K. N.: Analog nasalization of vowels. J. Speec ders *21:* 218–232 (1956).

Huffman, F.: The register problem i Khmer languages; in Jenner, Th

-.: An electromyographic study of velo-
eal function in speech. J. Speech Hear.
225–240 (1976).

F.; Baer, T.; Harris, K.; Niimi, S.:
ilatory effects of vowel quality on velar
. Phonetica *36:* 187–193 (1979).

G. N.: On the representation of vowel
Cornell University, Ithaca, unpublished,

I.: Patterns of soft palate movement in
iages. J. Phonet. *4:* 337–351 (1976).

: Phonetic and phonological rules of na-
n. UCLA Working Papers Phonet. 76

: Typology and universals of vowel sys-
Greenberg, Universals of human lan-
. II. Phonology, pp. 93–152 (Stanford
ty Press, Stanford 1978).

: Kluender, K. R.: The objects of speech
on. Ecological Psychol. *1:* 121–144

Walsh, M. A.: An auditory basis for the
-length effect in the perception of stops
es. J. acoust. Soc. Am. *85:* 2154–2164

-.: Vowel quality: the relation between
l and language-specific factors. UCLA
Papers Phonet. *58* (1983).

-.: Insights on vowel spacing; in Mad-
Patterns of sounds, chapter 9, pp. 136–
mbridge University Press, Cambridge

F.: On describing vowel quality; in
Jaeger, Experimental phonology, pp.
cademic Press, Orlando 1986).

M.; Mofokeng, S. M.: Textbook of
Sotho grammar (Longman, London

A.: An event approach to the study of
erception from a direct realist perspec-
honet. *14:* 3–28 (1986).

A.: Sound-producing sources as the ob-
perception: rate normalization and non-
perception. J. acoust. Soc. Am. *88:*
19 (1990).

A.; Smith, M. R.: Speech perception as
nalysis': an approach to the problems of
ce and segmentation; in Perkell, Klatt,
ce and variability in speech processes,
139 (Erlbaum, Hillsdale 1986).

Fox, R. A.; Trudeau, M. D.: A multidimensional scaling study of esophageal vowels. Phonetica *45:* 30–42 (1988).

Gandour, J.; Weinberg, B.: On the relationship between vowel height and fundamental frequency: evidence from esophageal speech. Phonetica *37:* 344–354 (1980).

Garner, W. R.: The processing of information and structure (Erlbaum, Potomac 1974).

Ghazeli, S.: Back consonants and backing coarticulation in Arabic; PhD diss. University of Texas at Austin (1977).

Grau, J. W.; Kemler Nelson, D. G.: The distinction between integral and separable dimensions: evidence for the integrality of pitch and loudness. J. exp. Psychol. general *117:* 347–370 (1988).

Gregerson, K.: Tongue-root and register in Mon-Khmer; in Jenner, Thompson, Starosta, Austroasiatic Studies, part 1. Oceanic Linguistics Special Publication 13, pp. 323–369 (University Press of Hawaii, Honolulu 1976).

Hall, B.; Hall, R.; Pam, M.; Myers A.; Antell, S.; Cherono, G.: African vowel harmony systems from the vantage point of Kalenjin. Afrika Übersee *57:* 241–267 (1974).

Hayes, B.: Diphthongization and coindexing. Phonology *7:* 31–71 (1990).

Hombert, J.-M.: Development of tones from vowel height? J. Phonet. *5:* 9–16 (1977 a).

Hombert, J.-M.: Consonant types, vowel height, and tone in Yoruba. Stud. Afr. Ling. *8:* 173–190 (1977 b).

Honda, K.: Relationship between pitch control and vowel articulation; in Bless, Abbs, Vocal fold physiology: contemporary research and clinical issues, pp. 286–297 (College-Hill Press, San Diego 1987).

Honda, K.; Fujimura, O.: Phonological vs. biological explanation – intrinsic vowel pitch and phrasal declination, 6th Vocal Fold Physiol. Conf., Stockholm 1989.

House, A. S.; Fairbanks, G.: The influence of consonant environment upon the secondary acoustic characteristics of vowels. J. acoust. Soc. Am. *25:* 105–113 (1953).

House, A. S.; Stevens, K. N.: Analog studies of the nasalization of vowels. J. Speech Hear. Disorders *21:* 218–232 (1956).

Huffman, F.: The register problem in fifteen Mon-Khmer languages; in Jenner, Thompson, Starosta, Austroasiatic studies, part 1. Oceanic Linguistics Special Publication 13, pp. 575–589 (University Press of Hawaii, Honolulu 1976).

Hyman, L.: Underspecification and vowel height transfer in Esimbi. Phonology *5:* 255–273 (1988).

Jackson, M. T. T.: Phonetic theory and cross-linguistic variation in vowel articulation. UCLA Working Papers Phonet. *71* (1988).

Keating, P. A.: Universal phonetics and the organization of grammars; in Fromkin, Phonetic linguistics. Essays in honor of Peter Ladefoged, pp. 115–132 (Academic Press, Orlando 1985).

Keating, P. A.: Underspecification in phonetics. Phonology *5:* 275–292 (1988).

Keating, P. A.: The window model of coarticulation: Articulatory evidence; in Kingston, Beckman, Papers in laboratory phonology, I: Between the grammar and physics of speech, pp. 451–470 (Cambridge University Press, Cambridge 1990).

Khabanyane, E. K.: Toward resolving disturbing misconceptions in learning of vowels in Southern Sotho (A language of South Africa). M. A. thesis. Cornell University (1989).

Kluender, K. R.; Diehl, R. L.; Wright, B. A.: Vowel length differences before voiced and voiceless consonants: an auditory explanation. J. Phonet. *16:* 153–169 (1988).

Krakow, R. A.; Beddor, P. S.; Goldstein, L.; Fowler, C. A.: Coarticulatory influences on the perceived height of vowels. J. acoust. Soc. Am. *83:* 1146–1158 (1988).

Kuehn, D. P.: A cineradiographic investigation of velar movement variables in two normals. Cleft Palate J. *13:* 88–103 (1976).

Kuenzel, H. J.: Photoelektrische Untersuchungen zur Velumhöhe bei Vokalen: erste Anwendungen des Velographen. Phonetica *34:* 352–370 (1977).

Ladd, D. R.; Silverman, K. E. A.: Intrinsic pitch of vowels in connected speech. Phonetica *41:* 31–40 (1984).

Ladefoged, P.: What are linguistic sounds made of? Language *56:* 485–502 (1980).

Ladefoged, P.: On dividing phonetics and phonology: Comments on the papers by Clements and by Browman and Goldstein; in Kingston, Beckman, Papers in laboratory phonology, I: Between the grammar and physics of speech, pp.

398–405 (Cambridge University Press, Cambridge 1990).

Ladefoged, P.: Vocal – a Macintosh computer program (unpublished, no date).

Lehiste, I.: Suprasegmentals (MIT Press, Cambridge 1970).

Lehiste, I.; Peterson, G. E.: Some basic considerations in the analysis of intonation. J. acoust. Soc. Am. *33:* 419–425 (1961).

Lindau, M.: Features for vowels. UCLA Working Papers Phonet. *30* (1975).

Lindau, M.: Vowel features. Language *54:* 541–563 (1978).

Lindau, M.: The feature 'expanded'. J. Phonet. *7:* 163–176 (1979).

Lindblom, B.: Vowel duration and a model of lip mandible coordination. Q. Prog. Status Rep., Speech Transm. Lab., R. Inst. Technol., Stockh., No. 4, pp. 1–29 (1967).

Lindblom, B.; Sundberg, J.: Acoustical consequences of lip, tongue, jaw, and larynx movement. J. acoust. Soc. Am. *50:* 1166–1179 (1971).

Linker, W.: Articulatory and acoustic correlates of labial activity in vowels: a cross-linguistic study. UCLA Working Papers Phonet. *56* (1982).

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

Melara, R. D.; Marks, L. E.: Perceptual primacy of dimensions: support for a model of dimensional interaction. J. exp. Psychol. hum. Perception Performance *16:* 398–414 (1990).

Moll, K. L.: Velopharyngeal closure on vowels. J. Speech Hear. Res. *5:* 30–37 (1962).

Moll, K. L.; Shriner, T.: Preliminary investigation of a new concept of velar activity during speech. Cleft Palate J. *4:* 58–69 (1967).

Ohala, J. J.: Experimental historical phonology; in Anderson, Jones, Historical linguistics. II. Theory and description in phonology, pp. 353–389 (North Holland, Amsterdam 1974).

Ohala, J. J.: Phonetic explanations for nasal sound patterns; in Ferguson, Hyman, Ohala, Nasalfest: Papers from a symposium on nasals and nasalization, pp. 289–316 (Language Universals Project, Stanford 1975).

Ohala, J. J.: The listener as a source of sound change. Papers from the parasession on language and behavior, pp. 178–203 (Chicago Linguistic Society, Chicago 1981).

Ohala, J. J.; Eukel, B.: Explaining the intrinsic pitch of vowels; in Channon, Shockey, In honour of Ilse Lehiste, pp. 207–215 (Foris, Dordrecht 1987).

Peterson, G. E.; Barney, H. L.: Control methods used in a study of vowels. J. acoust. Soc. Am. *24:* 175–184 (1952).

Pierrehumbert, J. B.: The phonology and phonetics of English intonation; PhD diss. Massachusetts Institute of Technology (1980).

Pierrehumbert, J. B.; Beckman, M. E.: Japanese tone structure (MIT Press, Cambridge 1988).

Riordan, C.: Control of vocal-tract length in speech. J. acoust. Soc. Am. *62:* 998–1002 (1977).

Ruhlen, M.: Nasal vowels; in Greenberg, Universals of human language, V. II. Phonology, pp. 203–242 (Stanford University Press, Stanford 1978).

Silverman, K. E. A.: The structure and processing of fundamental frequency contours; PhD diss. University of Cambridge (1987).

Smith, L. B.; Kemler, D. G.: Levels of experienced dimensionality in children and adults. Cognitive Psychol. *10:* 502–532 (1978).

Steele, S. A.: Interaction of vowel F_0 and prosody. Phonetica *43:* 92–105 (1986).

Stevens, K. N.; Fant, G.; Hawkins, S.: Some acoustical and perceptual correlates of nasal vowels; in Channon, Shockey, In honour of Ilse Lehiste, pp. 241–254 (Foris, Dordrecht 1987).

Stevens, K. N.; Keyser, S. J.; Kawasaki, H.: Toward a phonetic and phonological theory of redundant features; in Perkell, Klatt, Invariance and variability in speech processes, pp. 426–449 (Erlbaum, Hillsdale 1986).

Stevens, K. N.; Keyser, S. J.: Primary features and their enhancement in consonants. Language *65:* 81–106 (1987).

Stevens, K. N.: Some factors influencing the precision required for articulatory targets: comments on Keating's paper; in Kingston, Beckman, Papers in laboratory phonology, I: Between the grammar and physics of speech, pp. 471–475 (Cambridge University Press, Cambridge 1990).

Svantesson, J. O.: Tonogenetic mechanisms in Northern Mon-Khmer. Phonetica *46:* 60–79 (1989).

Syrdal, A.; Gopal, H.: A perceptual model of vowel recognition based on the auditory representation

---

of American English vowels. J.
79: 1086–1100 (1986).
Vaissiere, J.: Prediction of velum
phonological specifications.
122–139 (1988).
Vilkman, E.; Aaltonen, O.; Raimo
Oksanen, H.: Articulatory
changes vs. cricothyroid activity
intrinsic F_0 of vowels. J. Pho
(1989).
Westbury, J.; Keating, P.: Central
vowel duration (abstract). J. aco
S37 (1980).

; Eukel, B.: Explaining the intrinsic pitch
ls; in Channon, Shockey, In honour of
histe, pp. 207–215 (Foris, Dordrecht

i. E.: Barney, H. L.: Control methods
a study of vowels. J. acoust. Soc. Am. *24:*
(1952).

ert, J. B.: The phonology and phonetics
sh intonation; PhD diss. Massachusetts
of Technology (1980).

ert, J. B.; Beckman, M. E.: Japanese
icture (MIT Press, Cambridge 1988).

: Control of vocal-tract length in speech.
t. Soc. Am. *62:* 998–1002 (1977).

: Nasal vowels; in Greenberg, Univer-
uman language, V. II. Phonology, pp.
(Stanford University Press, Stanford

< E. A.: The structure and processing of
ental frequency contours; PhD diss. Uni-
f Cambridge (1987).

; Kemler, D. G.: Levels of experienced
onality in children and adults. Cognitive
10: 502–532 (1978).

.: Interaction of vowel F_0 and prosody.
a *43:* 92–105 (1986).

N.; Fant, G.; Hawkins, S.: Some acous-
perceptual correlates of nasal vowels;
ion, Shockey, In honour of Ilse Lehiste,
254 (Foris, Dordrecht 1987).

N.; Keyser, S. J.; Kawasaki, H.: Toward
tic and phonological theory of redun-
tures; in Perkell, Klatt, Invariance and
y in speech processes, pp. 426–449 (Erl-
illsdale 1986).

N.; Keyser, S. J.: Primary features and
ancement in consonants. Language *63:*
1987).

N.: Some factors influencing the preci-
ired for articulatory targets: comments
ing's paper; in Kingston, Beckman,
n laboratory phonology, I: Between
mar and physics of speech, pp. 471–
nbridge University Press, Cambridge

J. O.: Tonogenetic mechanisms in
Mon-Khmer. Phonetica *46:* 60–79

iopal, H.: A perceptual model of vowel
on based on the auditory representation

of American English vowels. J. acoust. Soc. Am. *79:* 1086–1100 (1986).

Vaissiere, J.: Prediction of velum movement from phonological specifications. Phonetica *45:* 122–139 (1988).

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

Westbury, J.; Keating, P.: Central representation of vowel duration (abstract). J. acoust. Soc. Am. *67:* S37 (1980).

Wright, J. T.: The behavior of vowels in the perceptual vowel space. University of California, Berkeley. Report of the Phonology Laboratory *5:* 127–163 (1980).

Zee, E.: Tone and vowel quality. J. Phonet. *8:* 247–258 (1980).

Prof. John Kingston
Department of Linguistics
South College
University of Massachusetts
Amherst, MA 01003 (USA)