568

Phonology Yearbook 3 (1986) 197–217
Printed in Great Britain

*Perceptual constraints and phonological change: a study of nasal vowel height**

Patrice Speeter Beddor
Rena Arens Krakow
Louis M. Goldstein
Yale University and Haskins Laboratories

ABSTRACT

To address the claim that listener misperceptions are a source of phonological shifts in nasal vowel height, the phonological, acoustic and perceptual effects of nasalisation on vowel height were examined. We show that the acoustic consequences of nasal coupling, while consistent with phonological patterns of nasal vowel raising and lowering, do not always influence perceived vowel height. The perceptual data suggest that nasalisation affects perceived vowel height only when nasalisation is phonetically inappropriate (e.g. insufficient or excessive nasal coupling) or phonologically inappropriate (e.g. no conditioning environment in a language without distinctive nasal vowels). It is argued that these conditions, rather than the inherent inability of the listener to distinguish the spectral effects of velic and tongue body gestures, lead to perceptual misinterpretations and potentially to sound change.

Phonologists have long supposed that listener misperceptions are a source of phonological change (e.g. Sweet 1888; Paul 1890; Durand 1956; Jonasson 1971; Ohala 1981). Listener misperceptions are presumably fostered by ambiguities in the acoustic signal with respect to articulation. That is, a given acoustic pattern may correspond (more or less closely) to more than one vocal tract configuration (e.g. [r] and [ʀ] are spectrally similar, but articulatorily very different). If a language learner were to identify incorrectly the articulatory source of an acoustic pattern (e.g. if [r] were perceived as [ʀ]), then, in attempting to imitate that pattern, the learner might produce the incorrectly reconstructed form rather than the original articulation. Thus the similarity of certain segments in the acoustic domain could lead to their reinterpretation in the articulatory domain (e.g. [r] reproduced as [ʀ]), and hence to sound change (e.g. /r/ > /ʀ/ in German; Jonasson 1971). (See Ohala 1981 for further discussion.)

The present study addresses the claim that listener misperceptions are a source of phonological change within the domain of nasal vowel height. Phonologically, there is substantial synchronic and diachronic evidence of raising and lowering of nasal vowels in languages of the world. It has been suggested that shifts in nasal vowel height originate with the listener, who attributes some of the complex acoustic consequences of nasal coupling to changes in tongue height, thereby perceiving nasal vowels as higher or lower than their non-nasal counterparts (Chen 1971; Ohala 1974; Wright 1980). As we will show below, this explanation for phonological shifts in vowel height is acoustically plausible, since some of the spectral consequences of coupling the nasal and oral tracts are similar to the effects of certain tongue body movements. However, this spectral similarity need not lead to *perceptual* confusion as to the articulatory source (i.e. tongue body *vs.* velic gesture) of the spectral pattern. In fact, as we will show, nasal coupling does not affect perceived vowel height when nasalisation of the vowel conforms to the phonetic and phonological structure of the listener's language. However, nasalisation does influence perceived vowel height under certain conditions which are inconsistent with that structure, as when a conditioning environment for vowel nasality is absent or nasal coupling is unexpectedly weak or strong. It is argued that these conditions, rather than the inherent inability of the listener to distinguish the spectral effects of velic and tongue body gestures, lead to perceptual misidentifications and potentially to sound change.

Our goal, then, is to shed some light on the extent to which phonological shifts in nasal vowel height can be attributed to listener misperceptions. We therefore consider three types of data: phonological (§ 1), acoustic (§ 2), and perceptual (§§ 3 and 4).

1 The phonological patterns

Diachronic and synchronic data from geographically distant and genetically unrelated languages indicate widespread phonological effects of nasalisation on vowel height. For example, in French, synchronic morphophonemic alternations attest to historical lowering of high and mid vowels and raising of low vowels, as in (1) (where N represents any nasal consonant):

(1)	[iN] ~ [æ̃]	fine/fin	'thin' (fem./masc.)
	[eN] ~ [æ̃]	plénitude/plein	'fullness/full'
	[yN] ~ [œ̃]	une/un	'one' (fem./masc.)
	[øN] ~ [œ̃]	jeûne/(à) jeun	'fast/fasting'
	[aN] ~ [ɑ̃]	planer/plan	'to glide/level'

Phonological studies comparing the height of contextual (allophonic) and non-contextual (phonemic or distinctive) nasal vowels to the height of corresponding oral vowels have found that, when differences occur, they are quite systematic across languages. Cross-language patterns of nasal vowel raising and lowering, based on Beddor (1982), are summarised in (2) (see Beddor 1982 for references). These patterns reflect synchronic allophonic and morphophonemic variation between oral and nasal vowel height in 75 languages, and are generally consistent with diachronic data and vowel inventory data from other cross-language surveys (Schourup 1973; Bhat 1975; Foley 1975; Ruhlen 1978):

(2) *Cross-language patterns of nasal vowel raising and lowering*

a. High (contextual and non-contextual) nasal vowels are lowered (e.g. nasalisation lowers /i/ and /u/ in Bengali, Ewe, Gadsup, Inuit and Swahili).
b. Low (contextual and non-contextual) nasal vowels are raised (e.g. nasalisation raises /a/ in Breton, Haida, Nama, Seneca and Zapotec).
c. Mid non-contextual nasal vowels are lowered (e.g. distinctive nasalisation lowers /e/ and /o/ in Maithili, Portuguese, Shiriana and Yuchi; distinctive nasalisation lowers /e/ (but not /o/) in Hindi, Mixtec and Kiowa Apache).
d. Mid back contextual nasal vowels are raised (e.g. /o/ or /ɔ/ is raised adjacent to N in Batak, Dutch and Nama).
e. A mid front contextual nasal vowel is raised in a language where the corresponding back vowel is also raised (e.g. /e/ and /o/ are raised adjacent to N in Irish, Basque and Havyaka Kannada); otherwise, mid front contextual nasal vowels are lowered (e.g. /e/ is lowered adjacent to N in Armenian, Campa, Fore and Tewa, but /o/ does not shift in these languages).

These patterns show that the phonological effects of nasalisation on vowel height involve the interaction of three factors: vowel height, vowel context, and vowel backness. Vowel height becomes centralised – that is, nasalisation lowers high vowels and raises low vowels. Vowel context (presence or absence of an adjacent nasal consonant) affects mid vowel height, and distinguishes lowering of mid non-contextual nasal vowels from raising of mid contextual nasal vowels. Vowel backness also primarily affects mid vowels, but a front–back asymmetry holds for all vowels, such that front vowels are more likely to be lowered than back vowels. More specifically, lowering of a back nasal vowel in a language implies lowering of the corresponding front nasal vowel in that language (Beddor 1982; see also Maddieson 1984).

2 Acoustic factors

The universality (in terms of genetic and geographic diversity) of these phonological patterns indicates that raising and lowering of nasal vowels are at least partially the result of phonetic constraints. Previously proposed phonetic explanations for shifts in nasal vowel height have appealed to articulatory (Pope 1934; Straka 1955; Lightner 1970; Pandey 1978), acoustic (Chen 1971; Ohala 1974; Wright 1980), and perceptual (Passy 1890; Haudricourt 1947; Martinet 1955; Ohala 1986) constraints. Indeed, a comprehensive explanation of the phonological data may well need to

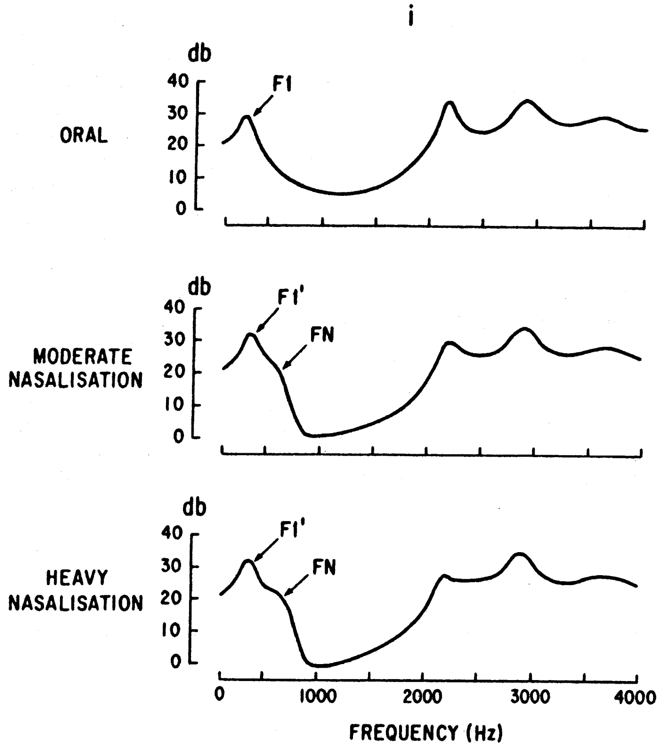

Figure 1

Vocal tract transfer functions for three versions of a high front vowel /i/ generated by articulatory synthesis: with no nasal coupling (top curve), with intermediate coupling (middle curve), and with large coupling (bottom curve). Nasal coupling shifted F1′ upward relative to F1 and introduced FN, which showed increased spectral prominence with larger coupling.

recognise the interaction of several phonetic, as well as non-phonetic, factors. However, we will consider here but a single phonetic factor, the effect of nasalisation on the first formant region of the vowel spectrum.

The main effect of vowel nasalisation is in the vicinity of the first formant. According to the acoustic theory of nasalisation, coupling of the nasal tract to the oral tract adds a pole–zero pair to the low-frequency region of the vowel spectrum (Fant 1960; Fujimura & Lindqvist 1971; Stevens *et al.* forthcoming). That is, the first formant F1 of the non-nasal vowel is replaced in the nasal vowel by a zero FZ and two formants, a shifted oral formant F1′ and a nasal formant FN. FN is almost cancelled by FZ when coupling magnitude is small, but becomes more and more prominent as coupling increases. F1′ typically differs in frequency, and

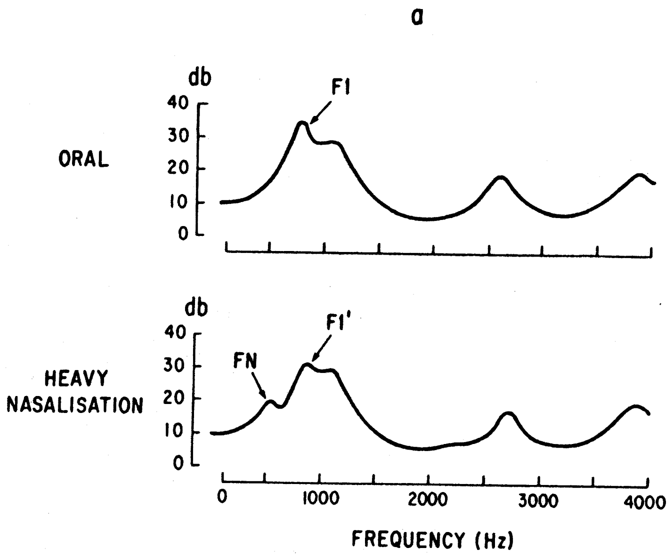

Figure 2

Transfer functions for oral [ɑ] (top curve) and nasal [ɑ̃] (bottom curve) generated by articulatory synthesis. Nasal coupling added a low-frequency FN and increased the frequency of F1′ relative to F1.

has a wide bandwidth and low amplitude, relative to F1 of the uncoupled oral tract (House & Stevens 1956; Mrayati 1975; Hawkins & Stevens 1985). Some of these spectral properties of nasal vowels are illustrated in Fig. 1 by the vocal tract transfer functions for oral and nasal versions of a high front vowel generated on the Haskins Laboratories articulatory synthesiser (described below). As velopharyngeal coupling is increased from no coupling for oral [i] (top curve) to intermediate coupling (middle curve) and large coupling (bottom curve) for nasal [ĩ], the frequency of F1′ shifts upwards and FN becomes increasingly prominent.

2.1 First formant frequency

Shifts in the frequency of F1′ in the nasal vowel relative to F1 in the oral vowel are of special interest here, since the frequency of F1 bears an inverse relation to vowel height. The acoustic theory of vowel nasalisation predicts that nasal coupling increases the frequency of the first oral formant, that is, F1′ > F1 (Fujimura & Lindqvist 1971; Mrayati 1975). This increase might lead us to expect nasalisation to lower perceived vowel height. However, this expectation ignores the fact that the upward-shifted F1′ is not necessarily the first peak in the nasal vowel spectrum. The lowest-frequency formant in the nasal vowel is located between F1 of the uncoupled oral tract and the lowest resonant frequency of the nasal tract

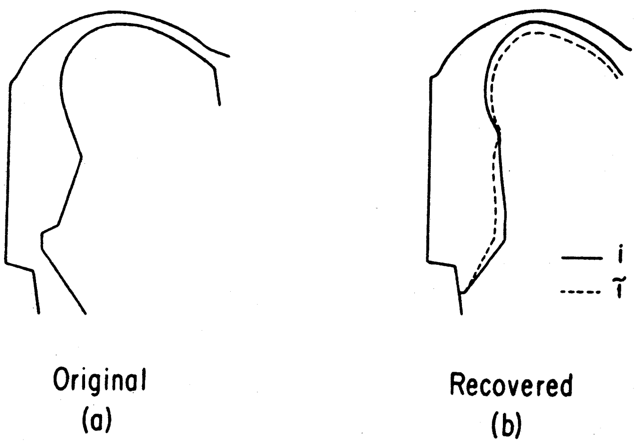

Figure 3

(a) Vocal tract shape of articulatory synthesiser used to compute the transfer functions in Fig. 1. (b) Vocal tract shapes recovered from the formant frequencies of the transfer functions in Fig. 1 (see text).

when closed at the coupling end (probably 200–400 Hz; Fujimura & Lindqvist 1971; Stevens *et al.* forthcoming). So when F1 of the oral vowel is relatively high (as in low vowels), the first formant of the coupled system is a low-frequency FN, as seen for low back [ɑ] and [ɑ̃] in Fig. 2. In contrast, the first formant of the high nasal vowel that was shown in Fig. 1 is the upward-shifted low-frequency oral formant. It follows that the frequency of the first spectral peak is higher in a nasal vowel than in the corresponding oral vowel when the vowel is high, but lower when the vowel is low. This is consistent with the centralising effect of nasalisation on phonological vowel height discussed above and thus provides a tentative acoustic explanation for high nasal vowel lowering and low nasal vowel raising (Wright 1980).

We can use a model of acoustic–articulatory relationships to demonstrate how these acoustic factors could lead to a sound change. The ability of a listener (or language learner) to reproduce an arbitrary speech sound must depend on knowledge that links the acoustic properties to their articulatory origins. If such knowledge were always perfect, then there would be no sound changes (for this reason, in any case) at all. Thus the knowledge brought to bear by the imitator is in some way imperfect (perhaps due to the inherent ambiguities mentioned earlier). As a model of an *extreme* case of such imperfection, let us imagine a listener (imitator) who has no knowledge of vowel nasalisation at all and who reproduces any vowel as oral. How will such a listener reproduce nasal vowels?

This question can be answered using the equations developed by Ladefoged *et al.* (1978) for calculating vocal tract shapes from formant

frequencies. These equations are based entirely on oral vowels. Thus the equations embody the acoustic–articulatory knowledge of a potential imitator ignorant of nasal vowels. We used these equations to calculate vocal tract shapes from the formants of the oral and heavily nasalised vowels shown in Fig. 1. For the nasal vowel, the shifted oral formant (F1′) was used as the lowest formant in the calculation. Fig. 3a shows the vocal tract shape (of the articulatory synthesiser) that was actually used to generate the transfer functions in Fig. 1 (except that the velar port was open in the nasal vowel). Fig. 3b shows the recovered vocal tract shapes using the Ladefoged *et al.* equations. Ignoring obvious differences in the pharynx (the equations do not recover the shape of the lower pharynx), the recovered [i] is very much like the original. However, the shape recovered for [ĩ] is substantially lower. Thus, lack of knowledge of the effects of nasalisation results in a high vowel being reproduced as a lower (oral) vowel. It is in this fashion that a sound change could develop. Of course, it is unlikely that any potential imitator has *no* knowledge of nasalisation – the model simply shows the degree of lowering that would be expected in the most extreme case.

2.2 Centre of gravity

Although the effects of nasal coupling on the location of the first peak in the vowel spectrum are consistent with contraction of the height dimension, they do not appear to account for the front–back asymmetry in the phonological data. If we extend our acoustic measure of oral and nasal vowels to include not only frequency of the first spectral peak, but frequency and relative amplitude of spectral peaks in the low-frequency region, we arrive at a more comprehensive explanation of the phonological patterns. Chistovich and her colleagues have found that perceived height of oral vowels reflects a 'centre of gravity' determined by the frequency and amplitude of spectral prominences in the F1–F2 region (Bedrov *et al.* 1978; Chistovich & Lublinskaya 1979; Chistovich *et al.* 1979). Due to the complex acoustic effects of nasal coupling, nasalisation can cause a shift in the centre of gravity of the vowel spectrum that need not correspond to a parallel shift in the frequency of the first spectral peak. For example, in the naturally produced mid front vowels in Fig. 4, the frequency of the first spectral peak is lower in nasal [ẽ] than in oral [e], but the overall effect of the pole–zero–pole combination in the low-frequency region of the nasal vowel is to pull up the centre of gravity relative to the oral vowel.[1]

Beddor (1982) measured the centre of gravity of oral and nasal vowel tokens from several languages by calculating the average frequency of the area under the spectral envelope in the F1–F2 region. This measure was consistently higher for [ĩ ẽ] than for [i e], lower for [æ̃ ã õ] than for [æ a o], and roughly the same for [ũ] and [u]. Assuming that an increase in centre of gravity lowers perceived vowel height and a decrease raises perceived height, we would expect nasalisation perceptually to lower /i e/, raise /æ a o/, and have little effect on /u/. Thus, oral–nasal differences

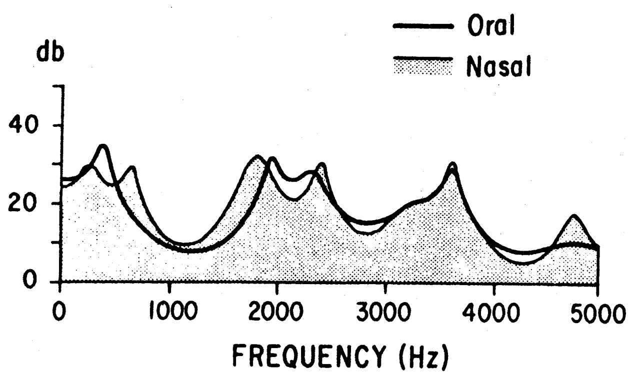

Figure 4
LPC spectra of oral [e] (unfilled) and nasal [ẽ] (filled) produced by a Hindi speaker. The nasal spectrum has a lower-frequency first peak, but a higher-frequency centre of gravity, than the oral spectrum.

in centre of gravity are consistent with the front–back asymmetry of the phonological data as well as high–low centralisation.

3 Perceptual validation

We have shown that the effects of nasal coupling on the low-frequency region of the vowel spectrum are generally consistent with the phonological patterns of nasal vowel raising and lowering. However, the acoustic data 'explain' the phonological shifts only if the listener is misled by the resemblance between spectral changes due to nasal coupling and those due to tongue body movements; that is, if the listener has imperfect knowledge of acoustic–articulatory relations, as discussed above. Rather than assign all of the spectral consequences of nasalisation to the velic gesture that couples the oral and nasal tracts, the listener must incorrectly attribute some of those spectral effects to a tongue gesture that modifies the oral tract configuration. Is there empirical evidence of such misperceptions? In answering this question, we hope to shed light not only on the role of listener misperceptions, but also on the relevance of context and speaker variability to vowel height shifts.

3.1 Perception of non-contextual nasal vowels

Several studies have investigated the perception of nasal vowel height. Wright (1980) produced natural oral and nasal vowels having the same tongue configuration, but differing in the position of the velum. All

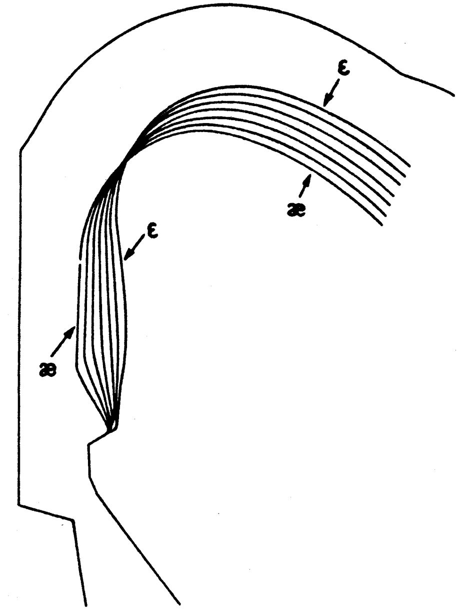

Figure 5
Vocal tract outlines of the seven steady-state vowel configurations specified by lowering and retracting the tongue body in equal articulatory steps from /ε/ to /æ/.

possible pairings of oral and nasal vowels were presented to listeners for similarity judgements. The perceptual vowel space constructed from listener responses showed centralisation of nasal vowel height relative to oral vowel height. Acoustic analysis of the vowels indicated that this centralisation did not always correlate with frequency differences in F1′ *vs.* F1, but might be partially due to the extra low-frequency FN in the nasal vowels.

In contrast to Wright's articulatorily matched vowels, Beddor (1984) paired oral and nasal vowels generated by formant synthesis. Listeners heard vowel sets in which a continuum of oral vowels (varying in the frequency of F1) was compared with a nasal vowel standard; they selected the oral vowel in each set which sounded most similar to the nasal standard. Listeners rarely chose the oral vowel in which F1 frequency was the same as F1′ frequency in the nasal vowel. In general, listeners' choices were pulled towards FN on the nasal vowel: when FN frequency was low, the oral vowel chosen as the 'best match' had a relatively low F1 frequency; when FN frequency was high, the oral match had a high F1.

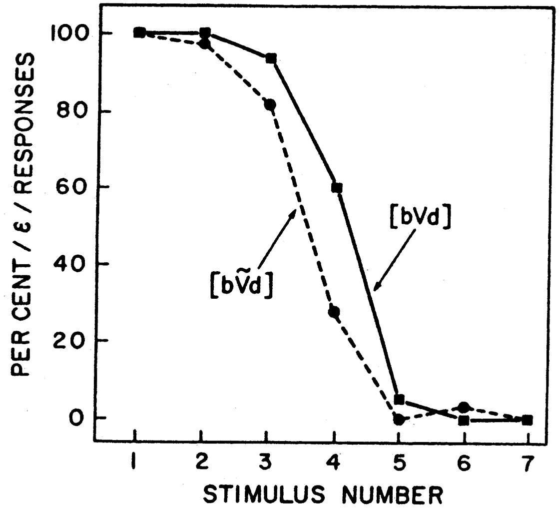

Figure 6

Pooled identification functions (n = 12) for the oral [bɛd–bæd] (squares) and the nasal [bɛ̃d–bæ̃d] (circles) continua.

Apparently (as in Wright's study), shifts in the spectral centre of gravity due to the added nasal formant affected perceived vowel quality.

In a recent study reported in Krakow *et al.* (in preparation), we used articulatory synthesis to investigate the effects of nasal coupling on perceived vowel height. The Haskins articulatory synthesiser allows specification of a mid-sagittal outline of the vocal tract by means of the positions of six articulatory parameters: jaw, hyoid, tongue body centre, tongue tip, lips, and velum. The program computes the area functions for the specified vocal tract outlines. Speech output is obtained after acoustic transfer functions are computed for these area values (see Abramson *et al.* 1981; Rubin *et al.* 1981).

In our study, we focused on the English /ɛ/–/æ/ contrast and generated seven vowels by systematically lowering and retracting the tongue body, as shown in Fig. 5. These vowel shapes were then embedded in an articulatory context appropriate for [b__d] and two 7-step continua were generated: oral [bɛd–bæd] and nasal [bɛ̃d–bæ̃d]. The only difference between the continua was that the velopharyngeal port was closed for the oral stimuli, but was opened 16·8 sq mm during the vowel portion of the nasal stimuli. (This size port opening yielded natural-sounding nasal vowels; see Krakow *et al.* for specification of the spectral properties of the stimuli.) Identification tapes for the two continua consisted of 10 tokens of each stimulus arranged in random order. Tapes were played to 12 phonetically naive native speakers of American English, who labelled the

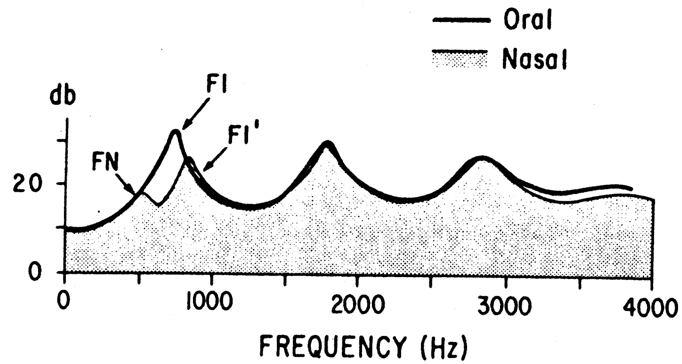

Figure 7

Transfer functions for the steady-state vowel portion of stimulus 4 from the oral (unfilled) and nasal (filled) continua. The first spectral peak has a lower frequency in the nasal vowel than in the oral vowel, but the predominant spectral peak in the nasal vowel is the upward-shifted F1′.

stimuli as *bed* or *bad*; they had no difficulty identifying the nasal vowel stimuli as such.

The identification functions in Fig. 6 show the percentage of /ɛ/ responses for both the oral (indicated by the squares) and the nasal (the circles) stimuli. There were fewer /ɛ/, and therefore more /æ/, responses to the nasal vowels than to the oral vowels; that is, nasalisation lowered perceived vowel height. This perceptual lowering is consistent with certain acoustic consequences of coupling the nasal tract to an /ɛ/-like oral tract configuration. For example, Fig. 7 gives the transfer functions for stimulus 4, which listeners more often labelled /ɛ/ when oral but /æ/ when nasal. Although FN and F1′ of the nasal vowel straddle F1 of the oral vowel, the predominant peak in the low frequencies of the nasal vowel spectrum is the upward-shifted F1′. The identification data can be interpreted as a tendency for listeners to associate the frequency shift induced by nasal coupling with lowering of the tongue body.

Our perceptual findings, like those of earlier studies, suggest that listeners have difficulty assessing the individual contributions of vowel quality and nasalisation to the spectral shape of the nasal vowel. Listeners may have attributed the spectral shifts in part to nasalisation, thus leading to the perception of a nasal vowel differing in height from the corresponding oral vowel. Alternatively, the spectral shifts may have been attributed entirely to oral tract shape, leading to the perception of a shifted oral vowel. Nonetheless, the data clearly show that spectral effects of nasalisation on vowels produced in isolation or in an oral context (i.e. non-contextual vowel nasalisation) are prone to misinterpretation by American listeners. And yet it would be premature to interpret these findings as evidence that listener misperceptions are a source of phonological shifts in nasal vowel height. These listeners may have been prompted to resolve the spectral effects of non-contextual nasalisation in terms of a tongue gesture only because of their unfamiliarity with distinctive nasal vowels.[2] That is, such misperceptions might not occur when the context for nasality is phonologically appropriate for the listener. To test this possibility, we need to look at the perception of non-contextual nasal vowels in languages with distinctive vowel nasalisation and also at the perception of contextual nasal vowels (i.e. nasal vowels in the immediate context of a nasal consonant) in languages with anticipatory or perseverative nasalisation. Some of our research addresses the second of these two issues.

3.2 Perception of contextual nasal vowels

Lowering of the velum for a nasal consonant has been found to begin during a preceding vowel to some degree in all languages investigated. Substantial anticipatory vowel nasalisation has been documented for many languages, including English (Malécot 1960; Moll 1962; Ali *et al.* 1971; Clumeck 1976). In Krakow *et al.* (in preparation), we tested our English-speaking subjects' perception of not only oral [bɛd–bæd] and non-contextual nasal [bɛ̃d–bæ̃d], but also contextual nasal [bɛ̃nd–bæ̃nd]. We speculated that in the [bṼnd] condition, the spectral effects of nasalisation on the vowel might be attributed to an anticipatory velic lowering gesture for the nasal consonant, thus allowing more accurate assessment of vowel configuration than in the [bṼd] condition.

Support for our speculation is provided by previous studies in which listeners were shown to be sensitive to coarticulatory information. In a study of vowel nasality, Kawasaki (1986) reported that perceived nasality of vowels in [mṼm] syllables was enhanced by attenuation of the adjacent nasal consonants. Her results suggest that listeners partially 'factored out' vowel nasalisation when the conditioning environment for nasalisation was perceptually salient. Ohala *et al.* (in preparation; see also Ohala 1981) looked at listeners' ability to recognise the coarticulatory fronting effects of apical consonants on adjacent /u/. They found that vowels ranging from [i] to [u] were more often labelled as back /u/ when flanked by apical consonants ([s__t]) than by labial consonants ([f__p]); that is, listeners apparently discounted some of the frontness of the vowels in the apical context as due to coarticulatory effects. Other studies have suggested that listeners are able to factor out coarticulatory effects not only of consonants on vowels, but also of vowels on consonants (e.g. Kunisaki & Fujisaki 1977; Mann & Repp 1980; Whalen 1981; Fowler 1984) and vowels on vowels (Fowler 1981). These data all suggest that knowledge of how phonetic units are coproduced influences speech perception. (More specific theoretical accounts of such facts have been proposed in Fowler 1983; Liberman & Mattingly 1985.) We thought that such knowledge might enable listeners to distinguish the effects of nasalisation from those of tongue shape on the spectrum of a contextual nasal vowel.

In our study, the contextual nasal condition [bɛ̃nd–bæ̃nd] was matched

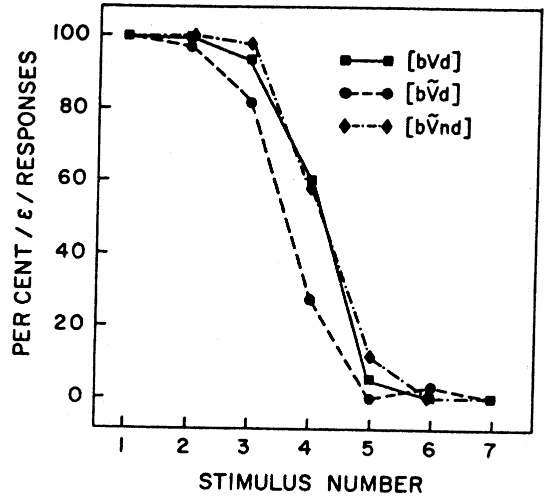

Figure 8

Pooled identification function ($n = 12$) for the contextual nasal [bɛ̃nd–bæ̃nd] continuum (diamonds) as compared with the oral [bɛd–bæd] and the non-contextual nasal [bɛ̃d–bæ̃d] functions (see Fig. 6).

as closely as possible to the oral and non-contextual nasal conditions described above. All vowel stimuli had the tongue shapes shown in Fig. 5. The contextual nasal continuum was the same as the non-contextual nasal continuum, except that the velopharyngeal port in the contextual nasal stimuli was open not only during the vowel, but remained open (at 16·8 sq mm) for 80 ms of the 137 ms alveolar occlusion, yielding natural-sounding [bVnd] sequences. Since the steady-state portions of corresponding contextual and non-contextual nasal vowels were identical, we hypothesised that if the perceived height of nasal vowels were strictly a function of their spectral characteristics, then listeners would judge vowel height to be the same in the two nasal conditions. However, if tacit knowledge of anticipatory nasalisation in English enabled listeners to factor out the spectral effects of contextual nasalisation, then perceived height of the contextual nasal vowels would be more, if not exactly, like that of the oral vowels.

Labelling responses to the contextual nasal stimuli were obtained from the 12 subjects who identified the oral and non-contextual nasal vowels. The experimental procedure was the same as that described above, except that subjects labelled the nasal stimuli as *bend* or *band* (as opposed to *bed* or *bad*). In Fig. 8, the identification responses to the contextual nasal [bṼnd] stimuli (the diamonds) are compared with the [bVd] and [bṼd] functions from Fig. 6. Notice that the point at which subjects shifted from /ɛ/ to

/æ/ responses (i.e. the 50% crossover point) in the [bṼnd] condition was the same as in the [bVd] condition; that is, contextual nasalisation had no effect on perceived vowel height.

3.3 Discussion

The perceptual data call into question simplistic accounts of the relation between listener misperceptions and nasal vowel height shifts. First, the fact that listeners did not misjudge nasal vowel height when provided with a conditioning environment for vowel nasalisation fails to support the idea that changes in *contextual* nasal vowel height are due to listeners misinterpreting the spectral effects of nasalisation as cues for vowel height. Second, although the finding that listeners misjudged nasal vowel height in the absence of a conditioning environment might appear to support listener misinterpretations as a source of *non-contextual* height shifts, this finding must be evaluated in light of the language background of the listeners. Our explanation for perceptual lowering of the non-contextual nasal vowels is that our American listeners did not expect a nasal vowel in the context [b__d] and consequently perceived the spectral changes introduced with nasal coupling as due at least in part to tongue configuration. This reasoning prompts us to expect different results if we were to obtain judgements of the non-contextual nasal vowels from listeners whose native language has distinctive vowel nasalisation. Since these listeners 'expect' nasal vowels to occur in oral (as well as nasal) contexts, we hypothesise that non-contextual nasalisation would have less of an effect – or perhaps no effect – on their perception of vowel height.

If listeners can separate the spectral effects of nasal coupling from those of tongue configuration, how then do we explain phonological raising and lowering of nasal vowels? We could, of course, turn to articulatory or even non-phonetic explanations, but the consistent correlations between the acoustic effects of nasalisation and the phonological patterns make us reluctant to reject an acoustic–perceptual approach. We can maintain that listener misperceptions lead to shifts in nasal vowel height if we can show that normal perceptual processing occasionally fails. Specifically, since listeners normally distinguish the acoustic consequences of velic *vs.* tongue body gestures, we need to show that this distinction can break down under certain conditions. In the next section, we consider two conditions which could lead to perceptual confusion and potentially to sound change.

4 Sources of perceptual confusion

4.1 Loss of conditioning environment

Ohala (1981, 1986) has argued that many sound changes in which loss of the conditioning environment co-occurs with the conditioned change can be explained by the listener's failure to detect the conditioning segment.

We believe a similar argument provides a tentative explanation for shifts in non-contextual nasal vowel height.

In the vast majority of languages that have distinctive nasal vowels, such vowels evolved from earlier sequences of phonemic oral vowels followed by nasal consonants (Ferguson 1963) or preceded by nasal consonants (Hyman 1972). One account of phonemicisation of vowel nasalisation with concomitant nasal consonant loss is that the perceptual salience of vowel nasality increased as the perceptual salience of the conditioning nasal consonant decreased (see Kawasaki 1986).[3] However, at the transition stage, distinctive vowel nasalisation is not fully integrated into the language. If listeners do not expect non-contextual nasal vowels but also do not perceive the now weakened nasal consonant, then they might attribute the acoustic effects of vowel nasalisation to either (a) nasal coupling, (b) change in tongue configuration, or (c) both nasal coupling and change in tongue configuration. Under these conditions, we would expect /VN/ or /NV/ to result historically in (a) /Ṽ/ with nasalisation but no height change, (b) /V′/ with height change but no nasalisation, or (c) /Ṽ′/ with height change and nasalisation.

Language data provide evidence of all three types of phonological change. There are numerous type (a) languages in which nasalisation has no marked effect on vowel height. For example, nasal vowel inventories were reportedly the same as oral vowel inventories in 77 of the 155 languages with phonemic nasal vowels surveyed by Ruhlen (1978). Possible examples of type (b) change include Greek **en > a* (**hekenton > hekaton*; Foley 1975), Colloquial Tamil final *eːn > æː* (Bhat 1975) and Old Norse, in which *i* and *u* lowered when a following nasal consonant was lost, but the nasality of the lowered vowels is uncertain (Bhat 1975). Type (c) languages are more difficult to identify, since distinctive nasalisation and height shift must be shown to have occurred more or less simultaneously. One such language appears to be French. Accounts of the evolution of French low non-contextual nasal vowels from non-low vowels followed by nasal consonants disagree on the relative order of distinctive nasalisation and vowel lowering (compare Pope 1934; Haden & Bell 1964; Martinet 1965; Entenman 1977), but the disagreement itself suggests that the two changes occupied roughly the same time period.

Evidence of type (a) languages (/VN/ > /Ṽ/) indicates that nasal consonant loss is not a sufficient condition for phonological shifts in nasal vowel height. At the same time, the existence of type (b) (/VN/ > /V′/) and type (c) (/VN/ > /Ṽ′/) languages suggests that nasal consonant loss is a possible trigger for such shifts. These phonological data correspond to our experimental results with American English speakers showing perceptual height shifts in [bṼd] sequences, although our results fail to distinguish whether listeners attributed all (as in type (b) languages) or only some (as in type (c) languages) of the spectral consequences of nasalisation to tongue height.

Our claim, then, is that listeners' ability to distinguish the acoustic consequences of velic *vs.* tongue body gestures might break down if the listener encounters a nasal vowel, but neither detects a conditioning nasal consonant nor expects non-contextual vowel nasalisation. We hypothesise that these conditions lead to ambiguity as to the nasality of the vowel. This uncertainty could in turn lead to changes in vowel height if the listener were to resolve at least some of the acoustic effects of nasalisation in terms of tongue configuration.

We have argued that, as a result of nasal consonant loss, there might be perceptual ambiguity leading to changes in vowel height. The next section postulates a second source of listener misperceptions which could influence the height of not only non-contextual, but also contextual, nasal vowels.

4.2 Variability in production

Most of our discussion of nasal vowels has approached vowel nasalisation as a binary distinction, such that vowels are either nasal or non-nasal. But there is considerable variation in degree of vowel nasalisation across vowel tokens, types and contexts, as well as across speakers and languages (Ohala 1971a; Ushijima & Sawashima 1972; Clumeck 1976; Benguerel *et al.* 1977; Henderson 1984). And as we have already seen (§2), different magnitudes of nasal coupling have different effects on the vowel spectrum.

What influence, then, might variability in degree of nasalisation have on vowel height? Consider, for example, a vowel followed by a nasal consonant. It seems reasonable to assume that the presence of the nasal consonant gives rise to certain expectations about the nasality of the vowel. If expectations are met, listeners should be able to factor the correct amount of nasalisation out of the vowel spectrum. But they might factor out too much (i.e. overcompensate) if nasalisation is unexpectedly weak, or too little (undercompensate) if nasalisation is excessive. (See Ohala 1981, 1986 for discussion of the possible role of overcompensation in sound change.) Both errors could affect perceived vowel height: overcompensation would reverse the direction of the height shifts predicted by acoustic factors, while undercompensation would yield the predicted shift.

Some of our [bɛ̃nd–bæ̃nd] results address this issue. The data presented in §4 were for a moderate (i.e. natural-sounding) amount of velopharyngeal port opening. But listeners were also tested on contextual and non-contextual nasal vowel stimuli produced with a small port opening (where nasalisation was judged by the experimenters as perceptually weak) and a large port opening (where nasalisation was judged as perceptually strong). The small port opening should raise the perceived height of the nasal vowels relative to the oral vowels if listeners overcompensate for weak nasalisation; the large opening should lower perceived height if listeners undercompensate for strong nasalisation.

Fig. 9 gives the identification responses to the contextual [bṼnd] and non-contextual [bṼd] stimuli with small (7·2 mm²) and large (24·0 mm²) port openings. In comparison with the oral [bVd] function, the nasal functions in Fig. 9a show that, although weak non-contextual nasalisation did not influence perceived vowel height, weak contextual nasalisation

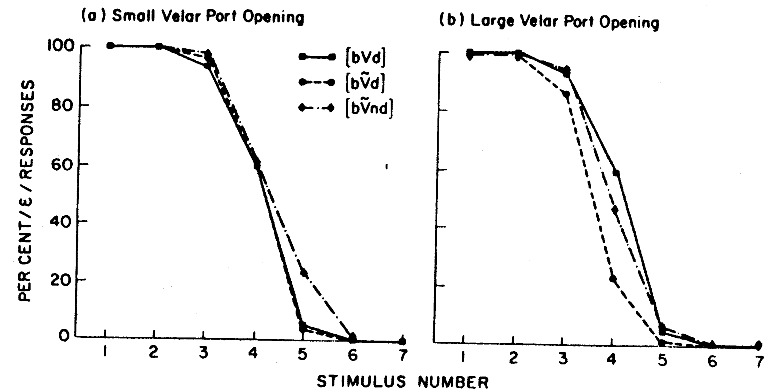

Figure 9

Pooled identification functions ($n = 12$) for the nasal continua generated with a small velar port opening (a) and a large port opening (b). The oral function is redrawn for comparison.

slightly raised perceived height (i.e. there were more /ɛ/ responses to the [bṼnd] stimuli than to either the [bVd] or the [bṼd] stimuli). Since our American English listeners presumably expected nasalisation in the contextual, but not the non-contextual, nasal conditions, our data suggest that listeners overcompensated for unexpectedly weak nasalisation. This finding lends support to the speculation by Ohala (1986) that phonological raising of mid contextual nasal vowels (as opposed to lowering of mid non-contextual nasal vowels) might be explained by listener overcompensation for contextual nasalisation.

In contrast, Fig. 9b shows that strong non-contextual and contextual nasalisation lowered perceived vowel height, with the non-contextual nasal vowels exhibiting greater lowering (i.e. the [bṼd] stimuli elicited the fewest /ɛ/ responses and the [bVd] stimuli the most). We interpret this result as evidence that listeners undercompensated for unexpectedly strong nasalisation.

The implication of our findings for sound change is that variability in degree of vowel nasalisation could cause perceptual uncertainty as to the relative contributions of the nasal and oral tracts to the vowel spectrum. Both weak and strong nasalisation could lead to height shifts because of listener failure to assess these contributions correctly.

5 Further questions and conclusion

Several issues concerning nasal vowel height have not yet been resolved. We have not yet studied the perception of nasal vowel height by speakers of a language with distinctive vowel nasalisation. While we have speculated that such listeners would show little or no effect of non-contextual nasalisation on perceived vowel height, absence of these data clearly limits our understanding of listeners' ability to factor out the effects of nasal coupling on the vowel spectrum. Unfortunately, this experiment may prove to be difficult to do, since many languages with distinctive nasal vowels show vowel quality differences between oral and nasal vowels, or phonotactic constraints against /CVNC/ sequences, or both (severely limiting the use of our stimuli for these purposes).

Another concern is that the timing of the velic gesture, like its magnitude (i.e. size of velopharyngeal opening), can differ in speakers' productions of nasal vowels, depending on the quality of the vowel, the speaker, and the language (Clumeck 1976). We still need to determine how temporal variability in the onset of the velic gesture affects the perceived height of nasal vowels. However, the present work leads us to conjecture that, for a given language, there is an 'expected' temporal pattern and that deviations from that pattern (e.g. premature velic lowering) would lead to perceptual ambiguity and perhaps phonological change in nasal vowel height.

In summary, we have seen that there are consistent cross-language phonological patterns of nasal vowel height defined by the interaction of vowel height, context and backness. We have also seen that a primary acoustic consequence of nasalisation is the introduction of a pole–zero pair in the vicinity of F1, the effect of which is to shift the centre of gravity in nasal vowel spectra relative to corresponding oral vowel spectra. These centre of gravity shifts can account for two important variables in the phonological data, vowel height and vowel backness, and therefore provide phonetic motivation for most of the phonological patterns if these acoustic effects of vowel nasalisation affect perceived vowel height. The perceptual data suggest that listeners misperceive nasal vowel height only when nasalisation is phonetically inappropriate (e.g. insufficient or excessive nasal coupling) or phonologically inappropriate (e.g. no conditioning environment in a language without distinctive nasal vowels). If inappropriate nasalisation were unique to the laboratory setting, then these perceptual findings would oblige us to reject the claim that listener misperceptions are a source of nasal vowel height shifts in natural languages. However, even though inappropriate nasalisation is not the 'norm', variations in degree of nasalisation and in the perceptual salience of the conditioning environment for vowel nasalisation are normal consequences of speech production and perception and as such are the raw material of nasal vowel height shifts. Thus we are brought, from another direction, to recognise the importance of variation in accounting for sound change (cf. Weinreich *et al.* 1968). It should be clear, however, that our acoustic–perceptual account of phonological changes in nasal vowel height has been restricted to the initiation of these changes. We have not attempted to specify the processes by which listener misconceptions become stable phonological patterns.

We have argued that listener familiarity with a particular phonetic and

phonological structure leads to certain expectations with respect to vowel nasalisation. Listeners correctly assess the contribution of nasal coupling to the vowel spectrum when these expectations are met, but when they are not, listeners apparently choose tongue configuration as an alternative source of the spectral effects of nasal coupling and thereby misperceive nasal vowel height. We conclude, then, that a comprehensive explanation of sound change in a language must take into account not only the physical (articulatory, acoustic or perceptual) origins of the change, but also the phonetic and phonological structure of the language, including variability in that structure.

NOTES

* This work was supported by NIH Grants HD-01994, HD-16591 and NS-07196. We thank Arthur Abramson, Björn Lindblom and John Ohala for helpful comments on earlier drafts of this manuscript and Carol Fowler for stimulating and encouraging the experimental work described here.

[1] The F2 differences in [e] and [ẽ] in Fig. 4 suggest that the two vowels may have been produced with different oral tract configurations; thus we cannot say to what extent the shift in centre of gravity is due to nasalisation *per se*.

[2] Although English does not typically have distinctive nasal vowels before voiced stops, an apparent exception to non-distinctive vowel nasalisation in American English occurs before voiceless stops. Malécot (1960) found that nasal consonants before voiceless stops are of extremely short duration, and may possibly be absent for some speakers, suggesting the existence of minimal pairs (e.g. *cat vs. can't*) differing only in vowel nasality.

[3] For some discussion of factors leading to weakening of nasal consonants, see Lightner (1970), Ohala (1971b), Schourup (1973), Foley (1975), Entenman (1977), Ruhlen (1978).

REFERENCES

Abramson, A. S., P. W. Nye, J. Henderson & C. W. Marshall (1981). Vowel height and the perception of consonantal nasality. *JASA* **70**. 329–339.

Ali, L., T. Gallagher, J. Goldstein & R. Daniloff (1971). Perception of coarticulated nasality. *JASA* **49**. 538–540.

Beddor, P. S. (1982). *Phonological and phonetic effects of nasalization on vowel height.* PhD dissertation, University of Minnesota. Reproduced by Indiana University Linguistics Club.

Beddor, P. S. (1984). Formant integration and the perception of nasal vowel height. *Haskins Laboratories Status Report on Speech Research* **SR-77/78**. 107–120.

Bedrov, Ya. A., L. A. Chistovich & R. L. Sheikin (1978). Frequency position of the 'center of gravity' of formants as a useful feature in vowel perception. *Soviet Physics Acoustics* **24**. 275–278.

Benguerel, A.-P., H. Hirose, M. Sawashima & T. Ushijima (1977). Velar coarticulation in French: a fiberscopic study. *JPh* **5**. 149–158.

Bhat, D. N. S. (1975). Two studies on nasalization. In Ferguson *et al.* (1975). 27–48.

Chen, M. (1971). Metarules and universal constraints on phonological theory. *Project on Linguistic Analysis, University of California, Berkeley* **13**. MC1–MC56.

Chistovich, L. A. & V. V. Lublinskaya (1979). The 'center of gravity' effect in vowel spectra and critical distance between the formants. *Hearing Research* **1**. 185–195.

Chistovich, L. A., R. L. Sheikin & V. V. Lublinskaya (1979). 'Centers of gravity' and spectral peaks as the determinants of vowel quality. In B. Lindblom & S. Öhman

(eds.) *Frontiers of speech communication research.* New York: Academic Press. 143–157.
Clumeck, H. (1976). Patterns of soft palate movements in six languages. *JPh* **4**. 337–351.
Durand, M. (1956). Du rôle de l'auditeur dans la formation des sons du langage. *Journal de Psychologie Normale et Pathologique* **52**. 347–355.
Entenman, G. (1977). *The development of nasal vowels.* PhD dissertation, University of Texas at Austin.
Fant, G. (1960). *Acoustic theory of speech production.* The Hague: Mouton.
Ferguson, C. A. (1963). Assumptions about nasals. In J. H. Greenberg (ed.) *Universals of language.* Cambridge, Mass.: MIT Press. 53–60.
Ferguson, C. A., L. M. Hyman & J. J. Ohala (eds.) (1975). *Nasálfest: papers from a symposium on nasals and nasalization.* Stanford: Language Universals Project.
Foley, J. (1975). Nasalization as universal phonological process. In Ferguson *et al.* (1975). 197–212.
Fowler, C. A. (1981). Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research* **24**. 127–139.
Fowler, C. A. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology: General* **112**. 386–412.
Fowler, C. A. (1984). Segmentation of coarticulated speech in perception. *Perception and Psychophysics* **36**. 359–368.
Fujimura, O. & J. Lindqvist (1971). Sweep-tone measurements of vocal-tract characteristics. *JASA* **49**. 541–558.
Haden, E. F. & E. A. Bell, Jr. (1964). Nasal vowel phonemes in French. *Lingua* **13**. 62–69.
Haudricourt, A. G. (1947). *En/an* en français. *Word* **3**. 39–47.
Hawkins, S. & K. N. Stevens (1985). Acoustic and perceptual correlates of the non-nasal–nasal distinction for vowels. *JASA* **77**. 1560–1575.
Henderson, J. B. (1984). *Velopharyngeal function in oral and nasal vowels: a cross-language study.* PhD dissertation, University of Connecticut.
House, A. S. & K. N. Stevens (1956). Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders* **21**. 218–232.
Hyman, L. M. (1972). Nasals and nasalization in Kwa. *Studies in African Linguistics* **3**. 167–205.
Jonasson, J. (1971). Perceptual similarity and articulatory reinterpretation as a source of phonological innovation. *Quarterly Progress and Status Report, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm* **1**. 30–42.
Kawasaki, H. (1986). Phonetic explanation for phonological universals: the case of distinctive vowel nasalization. In J. J. Ohala (ed.) *Experimental phonology.* New York: Academic Press. 81–103.
Krakow, R. A., P. S. Beddor, L. M. Goldstein & C. A. Fowler (in preparation). Coarticulatory influences on the perceived height of nasal vowels.
Kunisaki, O. & H. Fujisaki (1977). On the influence of context upon perception of voiceless fricative consonants. *Annual Bulletin, Research Institute for Logopedics and Phoniatrics, University of Tokyo* **11**. 85–91.
Ladefoged, P., R. Harshman, L. Goldstein & L. Rice (1978). Generating vocal tract shapes from formant frequencies. *JASA* **64**. 1027–1035.
Liberman, A. M. & I. G. Mattingly (1985). The motor theory of speech perception revised. *Cognition* **21**. 1–36.
Lightner, T. (1970). Why and how does vowel nasalization take place? *Papers in Linguistics* **2**. 179–226.
Maddieson, I. (1984). *Patterns of sounds.* New York: Cambridge University Press.

Malécot, A. (1960). Vowel nasality as a distinctive feature in American English. *Lg* **36**. 222–229.
Mann, V. A. & B. H. Repp (1980). Influence of vocalic context on perception of the [ʃ]–[s] distinction. *Perception and Psychophysics* **28**. 213–228.
Martinet, A. (1955). *Economie des changements phonétiques.* Berne: Francke.
Martinet, A. (1965). Les voyelles nasales du français. *Linguistique* **2**. 117–122.
Moll, K. L. (1962). Velopharyngeal closure on vowels. *Journal of Speech and Hearing Research* **5**. 30–37.
Mrayati, M. (1975). Etude des voyelles françaises. *Bulletin de l'Institut de Phonétique de Grenoble* **4**. 1–26.
Ohala, J. J. (1971a). Monitoring soft palate movements in speech. *Project on Linguistic Analysis, University of California, Berkeley* **13**. JO1–JO15.
Ohala, J. J. (1971b). The role of physiological and acoustic models in explaining the direction of sound change. *Project on Linguistic Analysis, University of California, Berkeley* **15**. 25–40.
Ohala, J. J. (1974). Experimental historical phonology. In J. M. Anderson & C. Jones (eds.) *Historical linguistics.* Vol. 2. Amsterdam: North-Holland. 353–389.
Ohala, J. J. (1981). The listener as a source of sound change. In C. S. Masek, R. A. Hendrick & M. F. Miller (eds.) *Papers from the parasession on language and behavior.* Chicago: Chicago Linguistic Society. 178–203.
Ohala, J. J. (1986). Phonological evidence for top-down processing in speech perception. In J. S. Perkell & D. H. Klatt (eds.) *Invariance and variability in speech processes.* Hillsdale, New Jersey: Lawrence Erlbaum Associates. 386–397.
Ohala, J. J., H. Kawasaki, C. Riordan & M. Caisse (in preparation). The influence of consonant environment upon the perception of vowel quality.
Pandey, P. (1978). A physiological note on the effect of nasalization on vowel height. *International Journal of Dravidian Linguistics* **7**. 217–222.
Passy, P. (1890). *Etude sur les changements phonétiques.* Paris: Librairie Firmin-Didot.
Paul, H. (1880). *Principien der Sprachgeschichte.* Halle: Niemeyer. Translated (1970) by H. A. Strong as *Principles of the history of language.* College Park, Maryland: McGrath.
Pope, M. K. (1934). *From Latin to Modern French.* Manchester: Manchester University Press.
Rubin, P., T. Baer & P. Mermelstein (1981). An articulatory synthesizer for perceptual research. *JASA* **70**. 321–328.
Ruhlen, M. (1978). Nasal vowels. In J. H. Greenberg (ed.) *Universals of human language.* Vol. 2: *Phonology.* Stanford: Stanford University Press. 203–241.
Schourup, L. (1973). A cross-language study of vowel nasalization. *Ohio State Working Papers in Linguistics* **15**. 190–221.
Stevens, K. N., G. Fant & S. Hawkins (forthcoming). Some acoustical and perceptual correlates of nasal vowels. In R. Channon & L. Shockey (eds.) *Festschrift for Ilse Lehiste.* Dordrecht: Foris.
Straka, G. (1955). Remarques sur les voyelles nasales, leur origine et leur évolution en français. *Revue de Linguistique Romane* **19**. 245–274.
Sweet, H. (1888). *History of English sounds.* Oxford: Oxford University Press.
Ushijima, T. & M. Sawashima (1972). Fiberscopic observation of velar movements in speech. *Annual Bulletin, Research Institute for Logopedics and Phoniatrics, University of Tokyo* **6**. 25–38.
Weinreich, U., W. Labov & M. Herzog (1968). Empirical foundations for a theory of language change. In W. Lehmann & Y. Malkiel (eds.) *Directions for historical linguistics.* Austin: University of Texas Press. 97–195.
Whalen, D. H. (1981). Effects of vocalic formant transitions and vowel quality on the English [s]–[š] boundary. *JASA* **69**. 275–282.
Wright, J. (1980). The behavior of nasalized vowels in the perceptual vowel space. *Report of the Phonology Laboratory, University of California, Berkeley* **5**. 127–163.