Perceptual Confusions and Phonological Change: How Confused is the Listener?

Author(s): Patrice Speeter Beddor and Rena Arens Krakow

Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetics and Phonological Universals (1998), pp. 320-334

Perceptual Confusions and Phonological Change: How Confused is the Listener?

Patrice Speeter Beddor and Rena Arens Krakow
University of Michigan / Temple University

1. Introduction

It has long been known that phonological systems are shaped, in part, by auditory factors. A crucial aspect of this shaping is the need to preserve perceptual distinctiveness among the contrastive speech sounds of a language. However, as is evident from daily communicative interactions, flawless productions of these speech sounds do not always occur and the perceptual mechanisms that enable listeners to maintain the distinctions are far from perfect. Hence, perceptual confusions arise more than occasionally. Because the communicative process is a dynamic one, with variations in the speech of a linguistic community having the potential to become a systematic property of that community's linguistic system, such confusions may lead to sound changes and consequently modifications of the set of phonological contrasts.

Acoustic and perceptual experimental evidence offers insight into the conditions leading to, and the consequences of, failure to maintain sufficient auditory distinctiveness in sound systems. Nasalization is a particularly interesting property to study in this regard. Perceptual evidence has shown, for example, that place distinctions in consonants are less salient when the consonants are nasal, than when they are oral (Mohr and Wang 1968, Hura et al. 1992). These distinctions are particularly non-salient in syllable-final position (Repp and Svastikula 1988, Ohala 1990, Beddor and Evans-Romaine 1995), which is compatible with articulatory evidence showing that nasal consonants have weaker oral constrictions in final than in initial position (Krakow 1989) and with acoustic evidence of a less abrupt change in spectral structure in VN than in NV syllables (Seitz et al. 1990, Manuel 1991). It is not only consonants that become more confusable when nasalized. For vowels, height differences are less distinct for nasal vowels than for their oral counterparts (Mohr and Wang 1968, Wright 1986) and perceived height may shift as a consequence of vowel nasalization (Krakow et al. 1988, Beddor and Hawkins 1990, Kingston 1991). The perceptual height effects are in keeping with acoustic evidence that nasal coupling, like tongue/jaw height, has its primary spectral effects on the first formant region of the vowel spectrum (Hawkins and Stevens 1985, Maeda 1993). Consequently, it is possible for the spectral information in that region to be ambiguous or misleading to the listener.

Clearly related to these experimental findings is an array of phonological facts. Perhaps most evident is the cross-linguistic frequency with which nasal (N) consonants in VNC clusters lose their distinct place constriction and take on that of the following oral (C) consonant. More relevant to the focus of the present paper is that, in most languages with distinctive vowel nasalization, these vowels have their origin in VN sequences in which the final consonant constriction disappeared while the coarticulatory effect of that nasal remained on the preceding vowel. That relatively non-salient place information in final nasal consonants contributes to consonant loss is supported by historical evidence that, prior to loss, place distinctions in final nasal consonants sometimes merge (Chen 1973, Tuttle 1991; cf. Hajek 1997). In addition, changes in vowel height due to vowel nasalization are not uncommon in the world's languages (Bhat 1975, Ruhlen 1978, Beddor 1982).

All of these phonological phenomena can be viewed as instances in which listeners fail to attribute a spectral property to the correct coarticulatory source.

2. A Complex Picture

But what might lead listeners to fail to attribute a spectral property to its coarticulatory source? Any attempt to link experimental evidence of relatively non-salient, and hence presumably unstable, perceptual distinctions with phonological change has to face the question of how an earlier stage of that sound system could tolerate the distinctions to begin with. In a now-classic paper, Ohala (1981) proposed that listeners are normally able to disambiguate at least certain types of perceptual similarities between speech sounds by applying 'reconstructive rules' that are based on listeners' language-specific knowledge of coarticulatory organization; sound changes may occur when listeners either fail to apply, or overapply, these rules. Under this view, normal perceptual processing would be particularly likely to break down if listeners failed to detect the coarticulatory source — that is, the conditioning environment — for the spectral property (Ohala 1981, 1986, 1993).

Conversely, listener confusions should be minimized when the acoustic signal provides clearly audible contextual information. Listeners who detect the source of the spectral property in question should be able to use their knowledge of coarticulatory structure to adjust for the variation and correctly identify the intended utterance. Perceptual studies that manipulate the consonantal context for vowel nasalization provide evidence that supports this approach. For example, Kawasaki (1986) found that attenuation of the nasal consonants in natural NVN utterances enhanced perceived vowel nasality. The American English listeners in that study were more likely to perceive a nasal vowel as oral when it was in a clearly audible nasal consonant context. This is consistent with the view that, when listeners detected the nasal consonant, they attributed the effects of nasalization on the vowel spectrum to the nasal consonant rather than to the vowel itself. Focusing on the height rather than on the nasality of nasal vowels, Krakow et al. (1988) found that American English listeners made accurate vowel height judgments of synthetic nasal vowels in an oral consonant context ([b_d]), but made vowel height mistakes on such vowels in a nasal consonant context ([b_nd]) (cf. Kingston and Macmillan 1995). One interpretation of this outcome is that listeners used the coarticulatory nasal context to attribute correctly the relevant aspects of the vowel's low-frequency spectral properties to nasality rather than tongue/jaw height.

More generally, experimental evidence (involving a variety of speech sounds and contexts), indicating that listeners attribute spectral effects of coarticulation to the context rather than the target segment, has led to the hypothesis that listeners adjust or 'compensate' for coarticulatory influences in identifying and discriminating speech sounds. We believe, consistent with Ohala's account and with our own account in previous work (e.g., Beddor et al. 1986, Krakow et al. 1988), that there is a link between failure to compensate for coarticulatory effects and perceptual confusions that give rise to sound change. At the same time, not all of the phonological and experimental data on nasal vowels are easily reconciled within this approach. Phonologically, we would expect that distinctive vowel nasalization would not occur in contexts in which the vowel is followed by a nasal consonant (because listeners should attribute the nasalization to context). Although such contrasts are rare, Hajek (1997) cites a number of languages with synchronic contrasts between [VN] and [ṼN]. We would also expect that contextual vowel nasalization would not give rise to vowel height shifts (again, because the context should enable listeners to attribute the spectral properties of the first formant region to nasalization rather than tongue height), but the phonological literature provides considerable cross-linguistic evidence of height shifts in contextually nasalized vowels (Beddor 1982). Experimentally, there is some evidence that the mechanism responsible for perceptual compensation for coarticulatory nasalization is more complex than that suggested by the data of Kawasaki (1986) and Krakow et al. (1988). In earlier work related to the current study (Krakow and Beddor 1991), we found that listeners asked to judge oral and nasal vowels in different contexts did not respond in a way indicating they consistently compensated for the coarticulatory effects of a nasal context. We expect that a better understanding of the types of adjustments listeners make — or fail to make —when presented with contextual information relevant to a perceptual decision will lead to a better understanding of the role of perceptual compensation in the structure of phonological systems. The two experiments described below were designed to tease apart some of these issues as related to vowel nasalization.

3. Rating Nasality

In the first experiment, listeners were asked to rate the relative nasality of vowels in a pair of syllables. The original stimuli were taken from a recording of a male native speaker of American English who produced multiple tokens of <u>bed</u>, <u>men</u>, <u>bode</u>, and <u>moan</u> in isolation. Two tokens of each of the 4 words, matched as closely as possible for vocalic pitch and loudness, were selected for use in the perceptual tests. Vowels were excised from their original contexts and were (a) spliced into the consonantal context of the other token of that word type (e.g. the vowel of <u>bed</u>$_1$ was spliced into the consonantal context of <u>bed</u>$_2$), (b) spliced into the consonantal context of the 'opposite' word type (e.g. the oral vowel of <u>bed</u>$_1$ was spliced into the nasal context of <u>men</u>$_1$), or (c) left excised from context (e.g. isolated <u>e</u>). (Excising did not affect the integrity of flanking nasal consonants: nasal consonant duration was sufficiently long that, even when vowel nasality was removed, [m] and [n] were unambiguously nasal.)

The splicing method yielded, for each vowel set (/ɛ/ or /o/), 2 tokens of each of 6 syllable types, as in Table 1. Within a given vowel set, all possible pairings of stimuli (including identity, included as a control condition) were created. To illustrate, pairs for the /ɛ/ set are shown in Table 2; the order of the pair members was counterbalanced in the actual test sequences. (For certain pair types in the /ɛ/ set, the duration of spliced vowels was also manipulated. Lax /ɛ/ — but not tense /o/ — was consistently shorter in nasal [m_n] than in oral [b_d] contexts. So that contextual appropriateness of vowel length would not be a factor in listeners' judgments, oral vowels were shortened to match nasal vowel duration in [mɛ̃n] - [mɛn] pairs and nasal vowels were lengthened to match oral vowel duration in [bɛd]-[bɛ̃d] and [ɛ]-[ɛ̃] pairs.

We presented the rating test to 16 native speakers of American English who were asked to decide for each pair which member had the more nasal vowel, or whether the pairs had equally nasal / non-nasal members. The test was blocked into /ɛ/ and /o/ sets, with half of the listeners responding first to the /ɛ/ set and the other half responding first to the /o/ set.

/ɛ/ set:	[bɛd]	[bɛ̃d]	[mɛn]	[mɛ̃n]	[ɛ]	[ɛ̃]
/o/ set:	[bod]	[bõd]	[mon]	[mõn]	[o]	[õ]

Table 1. Stimulus syllable types

Same-context pairs:		
[bɛd]-[bɛd]	[mɛ̃n]-[mɛ̃n]	[ɛ]-[ɛ]
[bɛd]-[bɛ̃d]	[mɛn]-[mɛ̃n]	[ɛ]-[ɛ̃]
[bɛ̃d]-[bɛ̃d]	[mɛn]-[mɛn]	[ɛ̃]-[ɛ̃]
Cross-context pairs:		
[bɛd]-[mɛ̃n]	[bɛd]-[ɛ]	[mɛ̃n]-[ɛ̃]
[bɛd]-[mɛn]	[bɛd]-[ɛ̃]	[mɛ̃n]-[ɛ]
[bɛ̃d]-[mɛn]	[bɛ̃d]-[ɛ̃]	[mɛn]-[ɛ̃]
[bɛ̃d]-[mɛ̃n]	[bɛ̃d]-[ɛ]	[mɛn]-[ɛ]

Table 2. Rating pairs for the /ɛ/ set

In presenting the results, we evaluate them relative to the *expected* results based on a compensation hypothesis which states that, in English (a language typically analyzed as having non-distinctive vowel nasalization), listeners attribute the effects of nasalization on the vowel spectrum to a flanking nasal consonant rather than to the vowel itself. The hypothesis predicts that listeners should make mistakes judging the nasality of pairs with vowels in nasal consonant contexts (except under the identity condition), but they should do relatively well judging vowel nasality in oral contexts and isolation. Fig. 1 gives the pooled percent correct responses of same-context pairs whose vowels differed in nasality, with /ɛ/ results in the left panel and /o/ results in the right. (Results for the control — i.e., identity — pairs are not presented, but averaged 83% correct across the 6 conditions.) Clearly, the general prediction is upheld: listeners' responses hover around 80% correct (n.b. close to control-condition accuracy) in the non-nasal contexts, but drop to around 40% correct in the nasal context. In addition, the /ɛ/ and /o/ sets elicited similar patterns of responses. For this reason, results for the two vowel sets are collapsed for the remainder of the comparisons.

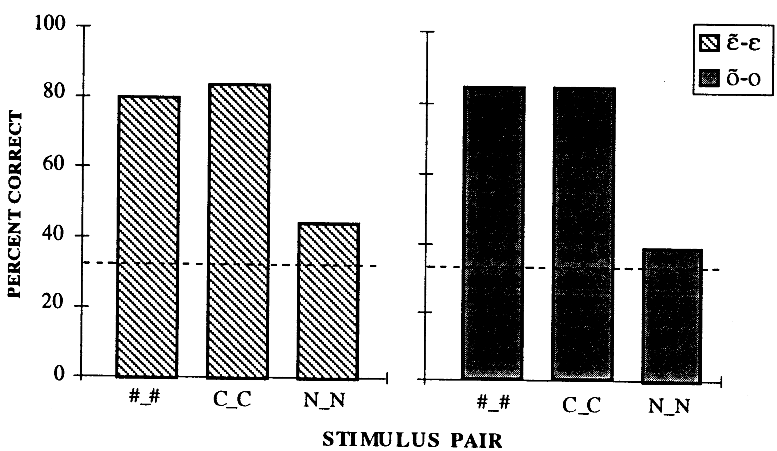

Fig. 1. Pooled percent correct rating responses to oral-nasal vowel pairs in identical contexts (striped bars for /ɛ/ and solid for /o/). Chance performance (horizontal dashed line) is 33%.

Rating pair	Correct response	Predicted (correct) response
[bɛd]-[ɛ]	'same'	'same'
[bɛ̃d]-[ɛ̃]	'same'	'same'
[bɛd]-[ɛ̃]	[ɛ̃]	[ɛ̃]
[bɛ̃d]-[ɛ]	[bɛ̃d]	[bɛ̃d]

Table 3. Predicted rating responses to cross-context comparisons involving non-nasal contexts

The cross-context comparisons provide additional insights into what occurs when vowels in pairs have the same, or different, nasality. The specific predictions for cross-context comparisons involving *non-nasal* contexts, where listeners should continue to perform accurately, are illustrated in Table 3 for the /ɛ/ set. The obtained results for these comparisons, shown in Fig. 2, support these predictions: as with the same-context comparisons, listeners' cross-context performance on vowels in non-nasal contexts remain at roughly 80% correct.

Turning to vowels in *nasal* contexts, the working hypothesis predicts that listeners should perform more poorly on cross-context comparisons in which one vowel is in a nasal context and the other is not. Because contextual information should lead listeners to compensate for nasality in the former case but not the latter, they should choose the wrong outcome for pairs involving [mɛ̃n] or [mɔ̃n]. We adopt, for the sake of exposition, a strong version of the compensation hypothesis which claims that American English listeners will attribute all of the nasalization on a vowel to a nasal context, if present. The resulting predictions are illustrated in Table 4 for [mɛ̃n]. The obtained results in Fig. 3 show that listeners have considerable difficulty judging the nasality of cross-context comparisons that involve [NṼN]: the near-chance performance on these nasal context pairs is a strikingly different pattern of results from that seen in Fig. 2 for non-nasal contexts.

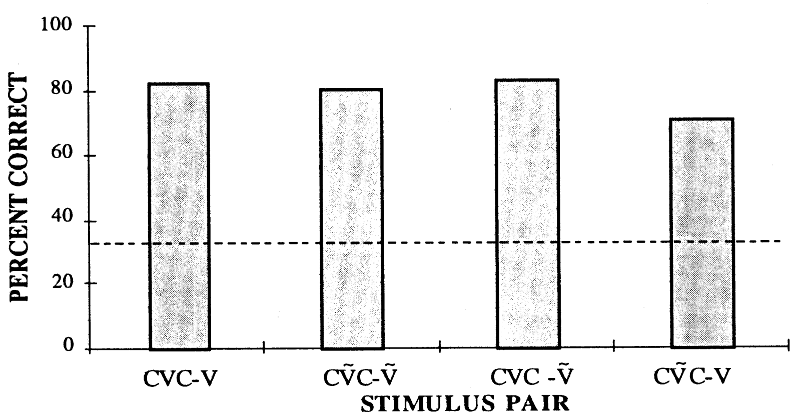

Fig. 2. Pooled percent correct rating responses to cross-context comparisons involving non-nasal contexts. Chance performance (horizontal dashed line) is 33%.

Rating pair	Correct response	Predicted (incorrect) response
[mε̃n]-[ε]	[mε̃n]	'same'
[mε̃n]-[bεd]	[mε̃n]	'same'
[mε̃n]-[ε̃]	'same'	[ε̃]
[mε̃n]-[bε̃d]	'same'	[bε̃d]

Table 4. Predicted rating responses to cross-context comparisons with [mε̃n]

What about *oral* vowels in *nasal* consonant contexts? Here we are less certain what to predict because we can imagine making a case for each of two conflicting hypotheses, both of which draw on the relation between production and perception. First, it seems reasonable to predict that listeners would accurately identify a vowel as oral in the coarticulatorily inappropriate context [N_N], much as we predicted them to be accurate at identifying a vowel as nasal in the coarticulatorily inappropriate context [C_C]. In such cases, lacking contextual support for presence ([CṼC]) or absence ([NVN]) of vowel nasalization is thought to render vowel nasality (or lack of nasality) perceptually obvious. Second, we might predict that listeners would be inaccurate at identifying a vowel as oral in a nasal context, parallel to the prediction that a nasal vowel will sound oral in a nasal consonant context (due to compensation for the contextual effect). Perhaps, for the [NVN] stimuli, listeners attempt to compensate for the expected effects of the nasal context, even though those effects are not present. Such compensation would make the vowel sound different than an oral vowel in an oral context or a nasal vowel in a nasal context (and the quality of such a vowel might be reported to be 'nasal' in the same way that individuals tend to describe the hyponasal speech of individuals with colds as 'nasal'). The first hypothesis predicts performance on pairs with [NVN] to be as accurate as performance on pairs with [CṼC]; the second predicts [NVN] performance to be as poor as [NṼN]. Interestingly, the results in Fig. 4 show that accuracy on pairs with [NVN] usually falls in between that of the other two types of pairs.

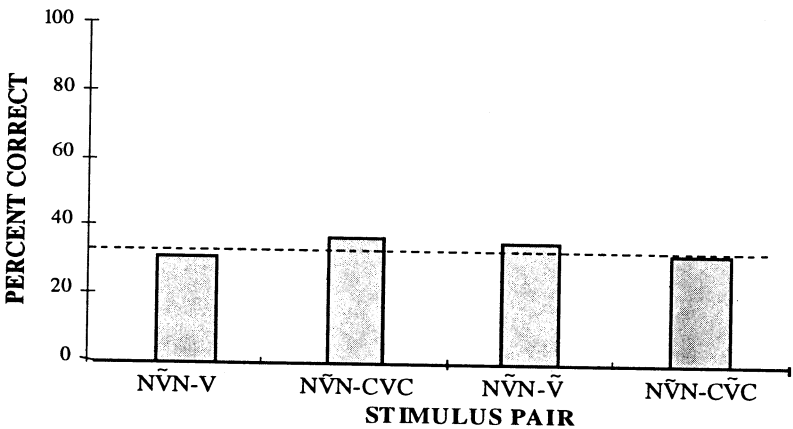

Fig. 3. Pooled percent correct rating responses to cross-context comparisons involving nasal vowels in nasal contexts. Chance performance (horizontal dashed line) is 33%.

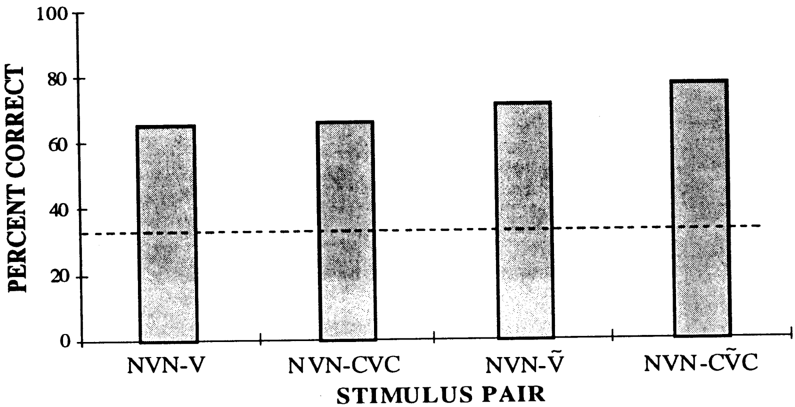

Fig. 4. Pooled percent correct rating responses to cross-context comparisons involving oral vowels in nasal contexts. Chance performance (horizontal dashed line) is 33%.

The general pattern of results is consistent with an account which suggests that listeners attribute vowel nasalization in a nasal consonant context to the context. Specifically, listeners are least likely to give the correct answer when pairs involve nasal vowels in nasal contexts (average: 34% correct). Listeners are increasingly 0likely to give correct answers in cross-context pairs as we progress from matches involving oral vowels in nasal contexts (average: 70% correct), to nasal vowels in oral contexts (average: 75% correct), to oral vowels in oral contexts (average: 84% correct).

However, a complex picture emerges when we consider the detailed pattern of responses shown by analysis not of percent correct responses, but rather of percent *choice* — that is, how often listeners chose each of the 3 options in a given pair type. Fig. 5 presents the 3-way response breakdown for the 4 cross-context pairs with a nasal vowel in a nasal consonant context from Fig. 3. Recall from the predictions above that the two left-most pairs, involving oral-to-nasal vowel pairings, were predicted to elicit 'same' responses. 'Same' was, indeed, the preferred choice for these two. On the other hand, Fig. 5 shows that when listeners chose one of the other options, they were more likely to make the correct nasality choice. That listeners were able to make this distinction suggests that the strong version of the compensation hypothesis is not correct.

The two right-most pairs in Fig. 5 are those [NṼN] pairs predicted not to elicit 'same' responses (i.e., nasal-to-nasal vowel pairs in which one pair member was contextually appropriate and the other was not). Although 'same' was, in fact, not listeners' most common choice, listeners did select the (correct) 'same' option approximately one-third of the time, indicating that compensation did not consistently lead them to make an incorrect nasality assessment.

Overall, performance on the rating task supports a modified version of the compensation hypothesis. When asked to judge the relative nasality of two vowels, listeners perform most poorly when one of the vowels is a nasal vowel in a nasal context, which is what we would expect if perceptual compensation leads listeners

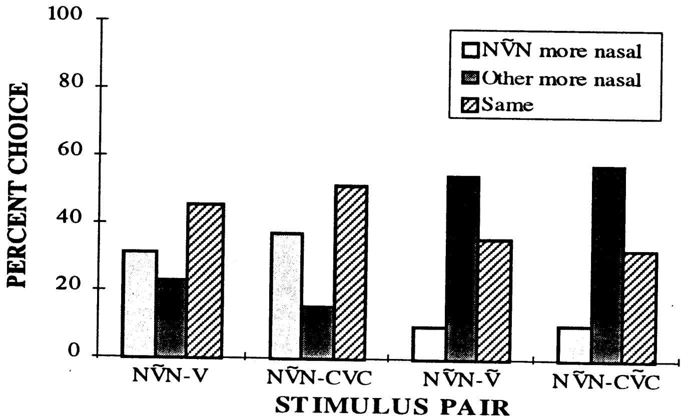

Fig. 5. Pooled percent choice responses to cross-context comparisons with nasal vowels in nasal contexts: frequency of listeners' choice of each of the 3 response options.

to attribute contextual vowel nasalization to context. At the same time, listeners do not respond in a way which indicates they usually hear nasal vowels in nasal contexts as oral, suggesting that perhaps the nasal context leads listeners to compensate *partially* for the effects of nasalization on the vowel spectrum. However, it may be premature to reach this conclusion solely on the basis of listeners' choices in a rating task, which requires listeners to make metalinguistic judgments of relative nasality. It is possible that the nature of the task interferes with our ability to determine the conditions under which oral and nasal vowels sound similar or different.

4. Discriminating Nasality

To address these issues further, we supplemented the rating task with a discrimination task which does not require listeners to make explicit nasality assessments. The same stimuli were arranged into a 4IAX discrimination paradigm in which each trial consisted of 2 stimulus pairs. In one pair, the vowels differed in nasality; in the other pair, the vowels had the same nasality and were, in fact, acoustically identical. (To preserve acoustic identity, the length manipulation applied to some stimulus pairs in the rating task was not used here.) To illustrate, discrimination trials for the /ɛ/ set are given in Table 5. As before, the tasks include the /ɛ/ and /o/ sets in separate blocks. (Note that not all possible trials were included: for trials involving vowels in consonantal contexts, all trials involved one pair in which vowel nasality was coarticulatorily appropriate. Consequently, we excluded trials such as [bɛ̃d]-[mɛ̃n]/[bɛ̃d]-[mɛn] or [bɛd]-[bɛ̃d]/ [bɛ̃d]-[bɛ̃d].)

The discrimination task was presented to the same 16 American English-speaking listeners who participated in the rating task. Listeners were told to select, for each trial, the pair whose vowels were more different. (Listeners took the discrimination test before the rating test, so as to avoid potential influence of nasality ratings on discrimination judgments.)

<u>Same-context trials</u>:		
[bɛd]-[bɛ̃d] / [bɛd]-[bɛd]	[ɛ]-[ɛ̃] / [ɛ]-[ɛ]	
[mɛn]-[mɛ̃n] / [mɛ̃n]-[mɛ̃n]	[ɛ]-[ɛ̃] / [ɛ̃]-[ɛ̃]	
<u>Cross-context trials</u>:		
[bɛd]-[ɛ̃] / [bɛd]-[ɛ]	[mɛn]-[ɛ̃] / [mɛn]-[ɛ]	[bɛd]-[mɛ̃n] / [bɛd]-[mɛn]
[bɛ̃d]-[ɛ̃] / [bɛd]-[ɛ]	[mɛn]-[ɛ̃] / [mɛ̃n]-[ɛ̃]	[bɛd]-[mɛ̃n] / [bɛ̃d]-[mɛ̃n]
[bɛd]-[ɛ̃] / [bɛ̃d]-[ɛ̃]	[mɛ̃n]-[ɛ] / [mɛn]-[ɛ]	
[bɛ̃d]-[ɛ] / [bɛ̃d]-[ɛ̃]	[mɛ̃n]-[ɛ] / [mɛ̃n]-[ɛ̃]	

Table 5. Discrimination trials for the /ɛ/ set

Applying the same working hypothesis of perceptual compensation to this task again leads to the general prediction that listeners will perform more poorly on cross-context trials involving nasal vowels in nasal contexts than on other types of trials. (Accurate performance was expected, and obtained, on all same-context trials: mean performance on each of the 4 types of same-context trials was above 95% correct.) As seen in the results presented in Fig. 6, the prediction of good performance on cross-context trials with vowels in non-nasal contexts was upheld, with performance ranging from 85-91% correct across the 4 conditions.

Predicting discrimination of vowels in nasal contexts, beyond the general expectation that such trials should be harder to discriminate than those represented in Fig. 6, is again straightforward only for those trials with [NṼN] that do not also include [NVN]. Given the possible groupings of our stimuli, this means that only 2 of the 6 trials with [N_N] yield clear predictions, as in Table 6.

The results for trials with [N_N] - isolated vowel pairings (corresponding to the first 4 rows of Table 6) are given in Fig. 7. Overall performance is poorer than that on the [C_C] - isolated vowel pairings in Fig. 6, dropping by an average of 16% and consistent with the general prediction. Comparing performance within the set of [N_N] - isolated vowel trials, Fig. 7 shows that the two trial types involving

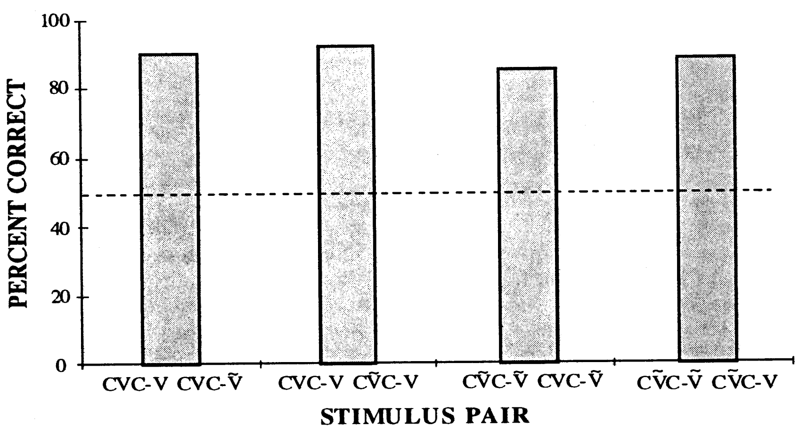

Fig. 6. Pooled percent correct discrimination responses to [C_C] - isolated vowel comparisons. Chance performance (horizontal dashed line) is 50%.

Trial type	Correct choice	Predicted choice
[mɛn]-[ɛ̃] / [mɛn]-[ɛ]	[mɛ̃n]-[ɛ]	?
[mɛn]-[ɛ̃] / [mɛ̃n]-[ɛ̃]	[mɛn]-[ɛ̃]	?
[mɛ̃n]-[ɛ] / [mɛn]-[ɛ]	[mɛ̃n]-[ɛ]	?
[mɛ̃n]-[ɛ] / [mɛ̃n]-[ɛ̃]	[mɛ̃n]-[ɛ]	[mɛ̃n]-[ɛ̃]
[bɛd]-[mɛ̃n] / [bɛd]-[mɛn]	[bɛd]-[mɛ̃n]	?
[bɛd]-[mɛ̃n] / [bɛ̃d]-[mɛ̃n]	[bɛd]-[mɛ̃n]	[bɛ̃d]-[mɛ̃n]

Table 6. Predicted discrimination choices for cross-context comparisons with [mɛ̃n]

[NṼN] - [V] pairs (the two right-most bars) were more difficult than trials not involving this comparison. It is interesting to note that this is the only pairing involving the [N_N] context and an isolated vowel in which the nasality of both vowels is contextually appropriate.

Turning our attention to discrimination of the consonantal context comparisons, we see in Fig. 8 that the contextually appropriate pairing of [NṼN] and [CVC] is like that of [NṼN] and [V] in that it hindered performance. However, the trials with [CVC] proved to be even more difficult. Discriminability of these trials was at chance regardless of whether the contextually appropriate vowels ([NṼN] - [CVC]) were compared with acoustically identical nasal ([NṼN] - [CṼC]) or oral ([NVN] - [CVC]) vowels.

5. Interpreting the Results

Overall, the rating and discrimination results exhibit similar response patterns. Vowel judgments in the two experimental paradigms showed the same relative order of difficulty for the different contextual conditions. Judgments of vowels in nasal contexts were more difficult (i.e., less accurate) than those of vowels in non-nasal contexts (oral contexts or isolation). Most difficult of all were judgments of nasal vowels in nasal contexts. Listeners responding to the rating task performed at

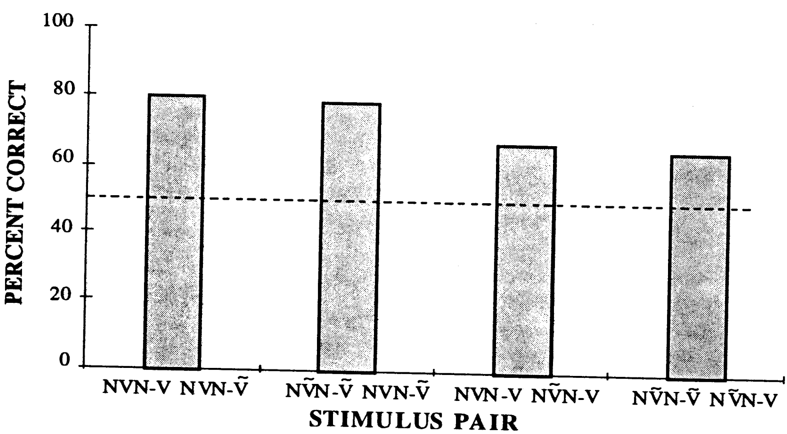

Fig. 7. Pooled percent correct discrimination responses to [N_N] - isolated vowel comparisons. Chance performance (horizontal dashed line) is 50%.

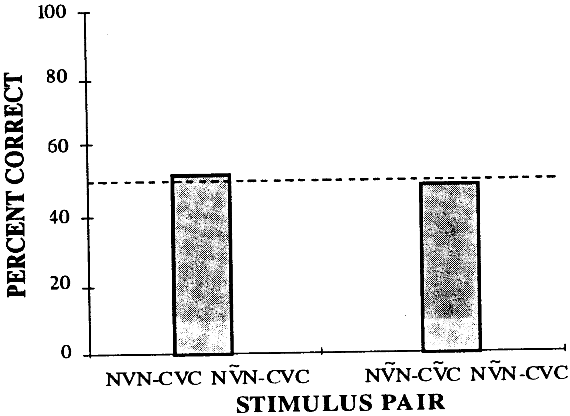

Fig. 8. Pooled percent correct discrimination responses to [N_N] - [C_C] comparisons. Chance performance (horizontal dashed line) is 50%.

chance level when judging the nasality of [NṼN] relative to [CVC], [V], and [Ṽ] (Fig. 3), and close to chance level relative to [NVN] (Fig. 1). Along similar lines, discrimination scores were at chance level for trials involving [NṼN] - [CVC] comparisons (Fig. 8) and were next lowest for [NṼN] - [V] comparisons (compare the right-most bars of Fig. 7 to the remainder of Fig. 7 and Fig. 6). These findings are generally consistent with a compensation hypothesis: when a target speech sound is in a coarticulatorily appropriate context, listeners' knowledge of coarticulatory organization lets them attribute the consequences of overlapping articulation on the target's spectrum to the coarticulatory source. Perceptual compensation is thought to *facilitate* speech perception under usual circumstances involving natural communicative settings. However, under laboratory conditions requiring listeners to make nasality judgments or to discriminate nasality, attributing spectral properties of the target to the context will give rise to listener *mistakes*: listeners attribute vowel nasality in a [NṼN] sequence to the nasal context and judge that vowel as less nasal than the same vowel in non-nasal contexts. This general outcome of the rating and discrimination tests is consistent with Kawasaki's (1986) findings using an experimental paradigm in which flanking nasal consonants were gradually attenuated and with our earlier (1991) findings using a matching paradigm.

Importantly, however, when listeners judge [Ṽ] as less nasal in nasal contexts than in oral contexts or isolation, this does not mean that they hear the [NṼN] vowel as *oral*. Indeed, the rating and discrimination results argue strongly against this interpretation. For example, in the rating task, if listeners heard contextually nasalized vowels as oral, they should have systematically chosen the incorrect option (in comparison with vowels in [C_C] context or isolation), but their performance was around chance rather than substantially below chance. In the discrimination task, if listeners heard the vowel in [NṼN] as oral, they should have systematically selected acoustically identical vowels in [NṼN] - [Ṽ] and [NṼN] - [CṼC] pairs as more different than non-identical but coarticulatorily appropriate vowels in

[NVN] - [V] and [NṼN] - [CVC] pairs. But here again, their performance was at or above chance.

Our findings clearly indicate that American English listeners do not - at least not consistently - hear contextually nasalized vowels as oral. At first blush, the data might seem to lend themselves to the interpretation that the chance and near-chance performance on some pairs involving [NṼN] stimuli (in both the rating and cross-context discrimination tasks) simply means that listeners are not certain about the nasality of such vowels and are therefore responding inconsistently. However, when we consider the entire dataset, we get a somewhat different view. For example, consider again the rating task comparisons involving [NṼN] paired with [CVC] or [V]. The detailed results (Fig. 5) showed that subjects selected [NṼN] as the 'more nasal' item considerably more frequently than they selected the oral alternative. Furthermore, performance on the discrimination task was around 70% correct for comparisons involving the contextually appropriate pair [NṼN] - [V]. Taken together, the results from the two experiments are most compatible with a view of partial compensation, that is, that listeners attribute some, but not all, of vowel nasalization in an [NṼN] sequence to the context.

6. Implications for Phonological Change

What are the implications of these findings for an account of phonological change which attributes (at least some of) the perceptual confusions leading to change to listeners' failure to compensate for coarticulatory effects? The account offered earlier in this paper (section 2), based largely on work by Ohala (1981, 1986, 1993 and others), predicts the occurrence of perceptual confusions when the conditioning environment for the spectral property at issue is not detected: in the absence of a coarticulatory source for that property, listeners analyze the property (e.g. nasalization) as an inherent characteristic of the target sound (e.g. vowel) rather than due to its contextual source (nasal consonant). On the other hand, clearly audible contextual information provides the requisite source; by enabling listeners to compensate perceptually for coarticulatory influences, source detection should block, or substantially reduce the likelihood of, confusion.

The perceptual data presented here point toward the need to modify this account. These findings indicate that unambiguous contextual information need not disambiguate the intended target sound. Indeed, listeners' inconsistency in judging the nasality of vowels in nasal contexts in some respects suggests that presence of the coarticulatory source leads to confusion rather than disambiguation. As just argued, however, patterns in the data indicate that listeners were not so confused in their responses as they were intermediate (nasal vowels in nasal contexts sounded less nasal than in oral contexts, but less oral than oral vowels). This pattern of partial compensation is not entirely surprising. If we assume that the listener weighs in all available information in making perceptual decisions (e.g. Hawkins 1995), then contextual information is but one factor in these decisions. Sometimes this information is not enough to disambiguate the acoustic signal. This may be particularly true of a property like contextual nasalization in American English, which can vary as a function of a variety of non-segmental factors such as syllable position and stress (Schourup 1973; see also Krakow 1993) as well as speaking rate (Kent et al. 1974, Bell-Berti and Krakow 1991, Krakow submitted), and also exhibits idiolectal and dialectal variation. Furthermore, the extent of contextual nasalization can vary as a function of vowel height (Clumeck 1976; this finding is consistent with vowel height effects on velar height, see Bell-Berti 1993 for a review).

Viewed in this way, one role that perception should play in phonological change would involve not so much a breakdown or failure of normal perceptual processes, but rather a shift in the relative weighting of factors contributing to the linguistic percept. When contextual information enables listeners to compensate only partially for coarticulatory effects, the spectral property is perceived as belonging to both the context and the target sound. If, indeed, this hypothesis of partial compensation is correct, then perceptually motivated sound changes in which the conditioning environment is *not* lost or particularly weak should be expected. For example, partial compensation for vowel nasalization in nasal consonant contexts leads, in principle, to at least two predictions. One prediction is that, since our findings indicate that nasalization can be perceived as belonging to a vowel even when the final consonant is clearly evident to the listener, distinctive nasal vowels could emerge in VN sequences without concomitant loss of the final nasal. A second prediction is that, if vowels in nasal contexts are perceived as (partially) nasalized, listeners should sometimes fail to disambiguate fully the spectral contribution of nasalization and tongue/jaw position in both contextual and distinctive nasal vowels, fostering vowel height shifts in both.

Reconsidering the phonological information provided in sections 1 and 2 in light of these predictions, we find that the evidence is consistent with the second prediction. Although the magnitude of the height shifts appears, in some cases, to be greater in distinctive than in contextual nasal vowels, there is no obvious difference in the frequency of occurrence (Beddor 1982). The situation corresponding to the first prediction is relatively rare: although there are languages exhibiting [VN] - [ṼN] contrasts, the usual historical development for contrastive nasalization involves nasal consonant loss. We expect this is linked to the fact that nasalization before final nasal consonants tends to be heavier than after initial nasals and that final consonants, including nasals, tend to have weaker oral tract constrictions than initial consonants (Krakow 1989, submitted). These factors conspire to make a final nasal consonant less perceptible than an initial nasal, especially in normal speaking (i.e., casual speech) conditions in which consonant constrictions can be further weakened (Krakow submitted) and anticipatory velar lowering more extensive (Bell-Berti and Krakow 1991).

What is our answer to the question posed in the title of the paper: how confused is the listener? The finding that listeners may partially compensate for a spectral property on a target sound when provided with salient contextual information suggests that listeners may be both more and less uncertain of the articulatory events that gave rise to the acoustic signal than previously thought. Listeners are less uncertain in that they can perceive a non-contrastive property as present on a target sound even when the property's coarticulatory source is also perceived. On the other hand, listeners' detection of the coarticulatory source need not fully disambiguate the acoustic signal; that is, such detection is apparently not a sufficient condition for unambiguous perception. We believe that perceptual ambiguity or partial compensation of the kind observed here is due largely to the apparent difficulty of perceptually resolving the spatial magnitude or temporal extent of gestural overlap among segments - not an unexpected difficulty in view of the considerable range of possible variation in these aspects of coarticulation under normal speaking conditions.

Acknowledgments
The research reported here was supported in part by NSF Grant SBR 9319597 to Patrice Beddor and NIH Grant NS-131617 to Haskins Laboratories. We wish to thank James Harnsberger for assistance in data collection, and the University of Michigan phonetics-phonology discussion group and participants in the BLS 24 Parasession for helpful comments.

References

Beddor, P. S. 1982. Phonological and phonetic effects of nasalization on vowel height. Doctoral dissertation, University of Minnesota, Minneapolis. (Reproduced by Indiana University Linguistics Club, 1983).

----, and D. Evans-Romaine. 1995. Listener misperceptions as a source of phonological assimilations: a study of syllable-final nasals. Rivista di Linguistica 7:1.145-74.

----, and S. Hawkins. 1990. The influence of spectral prominence on perceived vowel quality. J. Acous. Soc. of Am. 87.2684-2704.

----, R. A. Krakow, and L. M. Goldstein. 1986. Perceptual constraints and phonological change: a study of nasal vowel height. Phonology Yearbook 3.197-217.

Bell-Berti, F. 1993. Understanding velic motor control: studies of segmental context. Nasals, nasalization, and the velum, ed. by M. K. Huffman and R. A. Krakow, 63-85. Orlando, FL: Academic Press.

----, and R. A. Krakow. 1991. Anticipatory velar lowering: A coproduction account. J. Acous. Soc. of Am. 90.113-23.

Bhat, D. 1975. Two studies on nasalization. Nasálfest: papers from a symposium on nasals and nasalization, ed. by C. Ferguson, L. Hyman, and J. J. Ohala, 27-48. Stanford: Stanford University.

Chen, M. 1973. Nasals and nasalization in the history of Chinese. Doctoral dissertation, University of California, Berkeley.

Clumeck, H. 1976. Patterns of soft palate movements. J. of Phonetics 4.337-51.

Hajek, J. 1997. Universals of sound change in nasalization. Boston: Blackwell.

Hawkins, S. and K. N. Stevens. 1985. Acoustic and perceptual correlates of the non-nasal—nasal distinction for vowels. J. Acous. Soc. of Am. 77.1560-75.

Hawkins, S. 1995. Arguments for a non-segmental view of speech perception. Proceedings of the 13th International Congress of Phonetic Sciences, vol. 3.18-25. Stockholm.

Hura, S. L., B. Lindblom, and R. L. Diehl. 1992. On the role of perception in shaping phonological assimilation rules. Language and Speech 35.59-72.

Kawasaki, H. 1986. Phonetic explanation for phonological universals: the case of distinctive vowel nasalization. Experimental phonology, ed. by J. J. Ohala and J. J. Jaeger, 81-103. Orlando, FL: Academic.

Kent, R. D., P.J. Carney, and L. R. Severeid. 1974. Velar movement and timing: evaluation of a model of binary control. J. of Speech and Hearing Research 17. 470-88.

Kingston, J. 1991. Integrating articulations in the perception of vowel height. Phonetica 48.149-79.

----, and N. A. Macmillan. 1995. Integrality of nasalization and F_1 in vowels in isolation and before oral and nasal consonants: a detection-theoretic application of the Garner paradigm. J. Acous. Soc. of Am. 97.1261-85.

Krakow, R. A. 1989. The articulatory organization of syllables: a kinematic analysis of labial and velar gestures. Doctoral dissertation, Yale University, New Haven, CT.

----. 1993. Nonsegmental influences on velum movement patterns: syllables, sentences, stress, and speaking rate. Nasals, nasalization, and the velum, ed. by M. Huffman and R. Krakow, 87-113. New York: Academic Press.

----. Submitted. Articulatory organization of syllables: a review.

----, and P. S. Beddor. 1991. Coarticulation and the perception of nasality. Proceedings of the 12th International Congress of Phonetic Sciences, vol. 5.38-41. Aix-en-Provence.

----, P. S. Beddor, L. M. Goldstein and C. A. Fowler. 1988. Coarticulatory influences on the perceived height of nasal vowels. J. Acous. Soc. of Am. 83.1146-58.

Maeda, S. 1993. Acoustics of vowel nasalization and articulatory shifts in French nasal vowels. Phonetics and phonology: nasals, nasalization and the velum, ed. by M. K. Huffman and R. A. Krakow, 147-67. Orlando, FL: Academic Press.

Manuel, S. Y. 1991. Some phonetic bases for the relative malleability of syllable-final versus syllable-initial consonants. Proceedings of the 12th International Congress of Phonetic Sciences, vol. 5.118-21. Aix-en-Provence.

Mohr, B. and W. S.-Y. Wang. 1968. Perceptual distance and the specification of phonological features. Phonetica 18.31-45.

Ohala, J. J. 1981. The listener as a source of sound change. Papers from the parasession on language and behavior, ed. by C. S. Masek, R. A. Hendrick, and M. F. Miller, 178-203. Chicago: Chicago Linguistic Society.

----. 1986. Phonological evidence for top-down processing in speech perception. Invariance and variability of speech processes, ed. by J. S. Perkell and D. H. Klatt, 386-401. Hillsdale, NJ: Erlbaum.

----. 1990. The phonetics and phonology of aspects of assimilation. Papers in Laboratory Phonology I: between the grammar and physics of speech, ed. by J. Kingston and M. Beckman, 258-75. Cambridge, UK: Cambridge University Press.

----. 1993. Coarticulation and phonology. Language and Speech 36.155-70.

Repp, B. H. and K. Svastikula. 1988. Perception of the [m]-[n] distinction in VC syllables. J. of the Acoustical Society of America 83.237-47.

Ruhlen, M. 1978. Nasal vowels. Universals of human language, vol. II: Phonology, ed. by J. Greenberg, C. Ferguson, and E. Moravscik, 203-42. Stanford: Stanford University Press.

Schourup, L. 1973. A cross-language study of vowel nasalization. Ohio State Working Papers in Linguistics 15.190-221.

Seitz, P., M. McCormick, I. Watson, and R. A. Bladon. 1990. Relational spectral features for place of articulation in nasal consonants. J. Acous. Soc. of Am. 87.351-58.

Tuttle, E. F. 1991. Nasalization in Northern Italy: syllabic constraints and strength scales as developmental parameters. Rivista di Linguistica 3.23-92.

Wright, J. T. 1986. The behavior of nasalized vowels in the perceptual vowel space. Experimental phonology, ed. by J. J. Ohala and J. J. Jaeger, 45-67. Orlando, FL: Academic Press.