

TIME-VARYING PROPERTIES OF CONTEXTUALLY NASALIZED VOWELS: ACOUSTICS AND PERCEPTION

Marie K. Huffman

Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

ABSTRACT

Studies of English have shown that contextual nasalization is perceptible [8], and is used in speech processing [10]. However, when measured at single points in time, the spectral effects of contextual nasalization can appear quite subtle [7]. This suggests that time-varying properties of contextually nasalized vowels may be important to the perception of nasalization. This paper reports on acoustic and perceptual studies of changes in F1 prominence over time in contextually nasalized vowels of English. The results indicate that degree of F1 prominence, and change in F1 prominence over time, affect perception of nasality on vowels.

1. INTRODUCTION

Much research has gone into trying to characterize the acoustic effects of nasal coupling, and to determine which of these acoustic properties is important for the perception of nasalization. The effort is complicated by the fact that the spectral effects of nasalization can be relatively subtle [6,7], and that the perception of nasalization is closely bound up with the perception of vowel height [1,2]. We believe that the phonetic characterization of nasalization must take into consideration the possible role that time-varying properties of these vowels may play in their perception. Several researchers have suggested that change over time may be important to perception of nasalization on vowels [9,4], but little systematic investigation has been done in this area. This paper reports on acoustic and

perceptual studies of change over time in contextually nasalized vowels in English. These studies focus on the degree of F1 prominence, a spectral property which is consistently affected by nasal coupling. Although changes in F1 amplitude are frequently observed with vowel nasalization [5,6], their contribution to the perception of nasalization has been given relatively little attention.

2. F1 PROMINENCE IN NATURAL STIMULI

Before running perceptual experiments, we conducted an acoustic study to determine how F1 prominence was affected by contextual nasalization in natural speech. For the acoustic study, two male speakers of American English were recorded producing words of the form bVC, where C was either /d/ or /n/, and V was one of seven vowel qualities: [i, I, ei, E, ae, o, u]. The data were digitized at 10,000 samples per second, and then DFT spectra were computed for the vowels at 10 msec intervals.

In the DFT spectra, we quantified F1 prominence by calculating A1-H1, the difference in relative amplitude of A1 (the harmonic in F1 with the highest amplitude) and the fundamental. Figure 1 illustrates how A1-H1 was measured, and shows how A1-H1 decreases as nasalization increases on a contextually nasalized vowel.

To examine changes in F1 prominence over time on contextually nasalized vowels, we plotted A1-H1 for each spectral frame during the vowel. To better judge which effects on F1 were

attributable to nasalization as opposed to other factors, we plotted the data for the nasalized vowels along with similar data for comparable oral vowels. Figure 2 shows some examples of these combined plots, for one speaker. The second speaker showed a similar pattern.

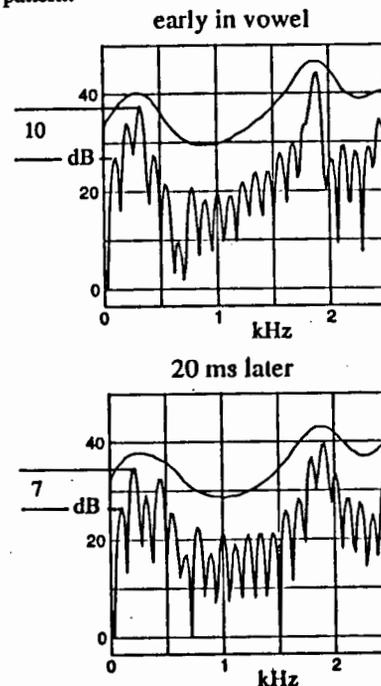


Figure 1. DFT spectra for two points in the vowel of "bin". An increase in nasalization over 20 msec results in a 3 dB decrease in A1-H1.

The data in Figure 2 show that, as we would expect, the contextually nasalized vowels had a smaller A1-H1 (less prominent F1), overall than the oral vowels. It should also be pointed out that A1-H1 values change over the course of the oral vowels as well as the nasalized vowels, suggesting that the prominence of F1 is affected by articulatory factors other than just nasal coupling. This underscores the importance of making relative, rather

than absolute, measures of the spectral effects of nasalization [3]. Finally, it should be noted that A1-H1 tends to decrease over the vowel, though this is true only for the latter half of [i]. On average, A1-H1 decreases by at least 4 or 5 db over the course of a contextually nasalized vowel. Given the patterns in the natural stimuli, we used synthetic speech to investigate the role of average F1 prominence, and change in F1 prominence over time, in the perception of nasalization on vowels.

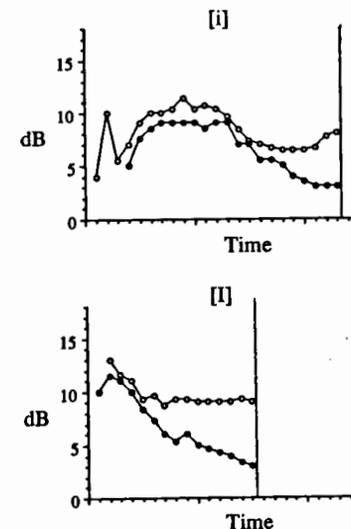


Figure 2. A1-H1 over time (10 msec intervals) for oral (open circles) and contextually nasalized vowels (filled circles), for one speaker of English.

3. PERCEPTION EXPERIMENT I: STATIC F1 PROMINENCE

The first experiment with synthetic speech focussed on how changes in average F1 prominence affect perceived nasalization. Stimuli were produced by starting with a synthesized oral vowel, and then decreasing F1 prominence by increasing F1 bandwidth. For a given vowel quality, several stimuli were produced, with different degrees of F1

prominence. For each item, F1 prominence was essentially constant over the duration of the vowel. The synthetic stimuli were chosen to have F1 prominence values which covered the range observed in natural speech items with the same vowel quality. Stimuli were constructed using 2 vowel qualities: [i], and [I]. Listeners heard a synthesized /bV/ syllable, followed by a vowel, and were instructed to choose which of two full words they felt the syllable could be an excerpt from. So, for example, on hearing [bi], the listener would circle either "bead" or "bean". 17 listeners participated in this experiment. The results of the tests are given in Figure 3, which shows the percentage of nasal responses for synthetic syllables having differing degrees of F1 prominence (as measured by A1-H1).

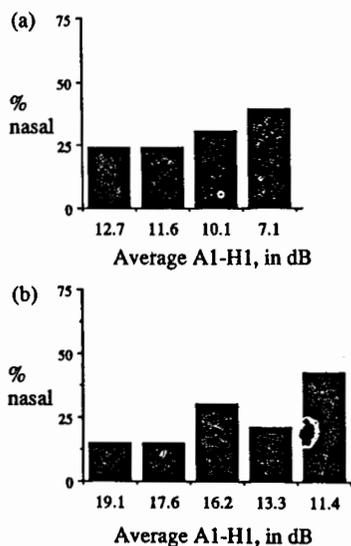


Figure 3. Percent nasal responses for synthetic /bi/ syllables (a) and /bI/ syllables (b), with static F1 prominence, at various values of A1-H1.

The majority of the time, the stimuli were heard as oral. However, the results suggest that decreased F1

prominence can contribute to production of a nasal percept. For stimuli made with the vowel [i], the percentage of times stimuli were heard as nasal increased with decreasing A1-H1 values, from about 25% nasal responses for a vowel with an A1-H1 typical of an oral vowel (12.7 dB), to about 40% nasal responses, for a vowel with an F1 that is about 5.5 dB lower. There is a similar pattern for [I], though the data appear to be a bit noisier.

4. PERCEPTION EXPERIMENT II: VARYING F1 PROMINENCE

We tested the effect of change in F1 prominence over time on perception of nasalization by comparing listener judgements of the nasality of stimulus pairs which were matched for vowel quality and overall average A1-H1, but which differed in having either an unchanging F1 prominence, as in the previous experiment, or a time-varying, decreasing, F1 prominence.

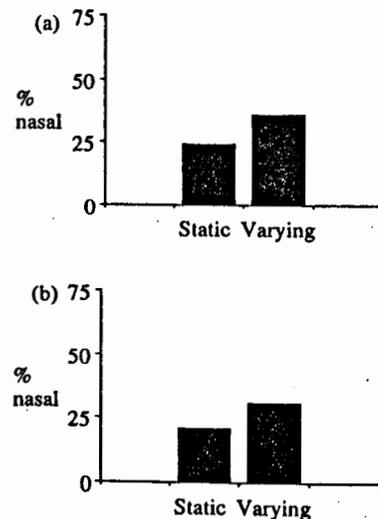


Figure 4. Percent nasal responses for stimuli with time-varying F1 prominence versus stimuli with static F1 prominence, matched for average A1-H1, for /i/ (a) and /I/ (b).

The time-varying vowel stimuli were synthesized with F1 prominence decreasing throughout the vowel. (The drop from vowel beginning to end was about 4 dB). Listeners heard synthetic /bV/ syllables containing these vowels within the same paradigm used in the previous experiment.

Figure 4 presents comparisons of the percentage of nasal responses for the stimuli with time-varying A1-H1 on the vowel, and their counterparts with static A1-H1 on the vowel. The time-varying stimuli show a higher percentage of nasal responses than the static stimuli, indicating that a decrease in F1 prominence over the course of the vowel results in more nasal responses than a simple, static reduction of F1 prominence.

5. SUMMARY

To conclude, we have seen that change in F1 prominence over time influences the perception of nasalization. By comparing perception of stimuli with static and time-varying F1 prominence, we determined that change over time is important, and that it is not just average F1 prominence which determines perceived nasalization. This is evidence for the importance of dynamic information in the perception of vowels. It also may have implications for predicting the likelihood of sound changes in which a contextually nasalized vowel becomes a contrastively nasalized vowel. Since physiological adjustments other than nasal coupling can affect F1 prominence, it is possible that changes in F1 prominence over time which come with diphthongization or laryngeal adjustments for voicing could contribute to a percept of nasalization. In combination with contextual nasalization, such effects could result in a stronger percept of nasalization, such that the language learner will be more inclined to posit a nasal vowel in that position, providing that other grammatical considerations do not prevent such an analysis. These questions await future research.

ACKNOWLEDGEMENTS

This research was supported by NIH grant #F32-NS08509. I am grateful to researchers at Haskins Laboratories and the MIT Speech Group for helpful discussions of this work.

REFERENCES

- [1] BEDDOR, P., KRAKOW, R. & GOLDSTEIN, L. (1986). "Perceptual constraints and phonological change: a study of nasal vowel height", *Phonology Yearbook*, 3, 197-218.
- [2] BEDDOR, P. & HAWKINS, S. (1990). "The influence of spectral prominence on perceived vowel quality", *J. Acoust. Soc. Am.*, 87, 2684-2704.
- [3] BELL-BERTI, F. (1980). "Velopharyngeal function: A spatial-temporal model", in *Speech and Language: Advances in Basic Research and Practice*, Vol. 4, 291-316.
- [4] BLADON, A. (1986). "Phonetics for Hearers", in G. McGregor, (ed.) *Language for Hearers*, Pergamon, Oxford.
- [5] DELATTRE, P. (1968). "Divergences entre nasalites vocalique et consonantique en francais", *Word*, 24, 64-72.
- [6] HOUSE, A. & STEVENS, K. (1956). "Analog studies of the nasalization of vowels", *J. Speech Hear. Dis.*, 21, 218-232.
- [7] HUFFMAN, M. (1990). "An acoustic study of the timing of contextual nasalization in English", *J. Acoust. Soc. Am. Vol. 87, Suppl. 1*, S67.
- [8] HUFFMAN, M. (1990). "The role of F1 amplitude in producing nasal percepts", *J. Acoust. Soc. Am. Vol. 88, Suppl. 1*, S54.
- [9] REENEN, P. T. VAN (1982). *Phonetic feature definitions*. Foris Publications, Holland.
- [10] WARREN, P. & MARSLER-WILSON, W. D. (1987). "Continuous uptake of acoustic cues in spoken word recognition", *Perception & Psychophysics*, 41, 262-275.