Coarticulation Effects on the Nasalization of Vowels Using Nasal/Voice Amplitude Ratio Instrumentation

PATRICA L. LARSON, M.A.
SANDRA L. HAMLET, PH.D.

Nasal coarticulation in phonetically controlled nonsense syllables was investigated in four normal adult speakers. Nasalization was determined using the ratio of a nasal accelerometer signal amplitude to airborn microphone signal amplitude. Measurements of nasalization were made at the midpoint of vowels and at a constant time from the nasal consonant. Nasal acoustical coupling was greater for high vowels than for low vowels in all consonant contexts. Nasalization was also greater for vowels between two nasal consonants than for vowels between a nasal consonant and a fricative or stop. Results for progressive versus regressive assimilation depended on the measurement strategy. For within-vowel measurements made a constant time from the nasal consonant, prenasal vowels showed greater nasalization than postnasal vowels. This nasal accelerometric technique shows promise for clinical assessment of articulatory details of velar function.

Noninvasive techniques for quantifying the effects of velar function in speech are of special interest for clinical applications such as biofeedback in rehabilitation and documenting the results of treatment. One promising method, using an accelerometer to detect nasalization, has been utilized by a number of investigators. In 1975, Stevens, Kalikow, and Willemain described the use of a miniature accelerometer to detect nasalization during speech. Subsequently, Stevens et al (1976) published data on peak nasal accelerometer signal measurements for vowels in normal and deaf speakers. They found that a small accelerometer signal could be received during non-nasalized vowel productions, but that vowels adjacent to nasal consonants had 11 to 20-dB larger peak accelerometer signal amplitudes than the non-nasalized vowels.

One problem with nasal accelerometry noted by Stevens et al (1976) and Garber and Moller (1979) was that it was sensitive to changes in vocal intensity. To reduce this limitation, Horii (1980) proposed an index of nasal coupling called the "Horii Oral Nasal Coupling Index (HONC)," which employed the ratio of nasal accelerometer signal amplitude to laryngeal accelerometer signal amplitude. Miniature accelerometers were placed on both the external neck and nasal skin surfaces. An alternative, substituting airborn voice signals picked up by a microphone for the laryngeal accelerometer signal was suggested in the later article (Horii and Lang, 1981). Similar systems for nasal vibration analysis were developed by Edgerton et al (1981) and Rendenbaugh and Reich (1985).

Using the HONC system, Horii (1983) investigated simulated hypernasality in normal speakers, finding a correlation coefficient of 0.92 between measured nasalization and judgments of the severity of hypernasality on a read paragraph. Redenbaugh and Reich (1985) and Reich and Redenbaugh (1985) also found high correlation coefficients (0.91 and 0.90) between listener judgments of naturally occurring nasality and accelerometric ratio scores, but only for sentences excluding nasal consonants. In English, coarticulated nasality on vowels is expected as an allophonic variant in proximity to nasal consonants, so coarticulated nasality was apparently not always judged as nasal speech, even though accelerometric scores reflected the presence of nasality.

Until recently the majority of research conducted using the ratio of nasal accelerometer signal to laryngeal accelerometer signal as a means of quantifying nasalization has investigated isolated sounds, or whole passages where an overall nasal-to-voice index was calculated.

Ms. Larson was Speech Pathologist at the Hitchcock Rehabilitation Center in Aiken, South Carolina. She is currently Speech Pathologist at St. Joseph's Hospital in Augusta, Georgia. Dr. Hamlet was Associate Professor at the University of Maryland, College Park, Maryland and is now Research Associate Professor in the Department of Otolaryngology at Wayne State University in Detroit, Michigan.

This study was supported in part by grant # DE-03631 from the National Institute of Dental Research.

Stone and Horii (1986) have extended such investigation to nasal coarticulatory effects, although still using the HONC index for whole passages of speech. They found that passages containing phonetic contexts designed to elicit regressive vs progressive assimilation did not differ significantly in HONC scores. Horii (1986), investigating temporal detail of nasal accelerometer signals, reported extended nasal assimilation on vowels for fast rates of speech.

The present study was designed to investigate whether phonetic contextual details of nasal coarticulation that have been found using other methodologies could also be demonstrated with nasal accelerometry. If so, this would extend the clinical usefulness of nasal accelerometric techniques to assessment of the dynamic articulatory aspects of velar function in speech.

METHOD

Subjects

The subjects were four female speakers of American English between the ages of 20 and 40 years. None had ever been identified as having a speech, language, or voice disorder.

Instrumentation

Nasal resonance was detected by placing a miniature accelerometer secured with double-sided adhesive tape approximately 5 mm from the nostril on the ala of one side of the nose (Lippmann, 1981). The accelerometer was one of the input options to a Visual Speech Training Aid (VISTA), which amplified, rectified, and smoothed the signal with a time constant of approximately 15 msec. The demodulated nasal accelerometer signal was tapped off of the VISTA internally just prior to its processing for video monitor display. The acoustic signal was transduced using a dynamic microphone mounted 6 inches from the lips. Both signals were recorded on a Hewlett-Packard 3960 Instrumentation Recorder.

For analysis, both signals were played back and the nasal accelerometer-to-voice microphone signal ratio (N/V) was obtained using an integrated circuit analog divider (Analog Devices 433). Before being input to the divider circuit, the microphone signal was also rectified and smoothed with a time constant of approximately 15 msec. Thus, both the accelerometer and microphone signals were processed in equivalent ways. The recorded microphone signal recorded accelerometer signal, and the N/V ratio signal were printed out simultaneously using a Honeywell 1858 Visicorder at a paper speed of 4 inches per second. Analog processing of signals was chosen as the most expedient means of instrumenting the study, not because of any inherent superiority over equivalent digital means.

Speech Sample

Twenty different nonsense syllables were chosen having the following phonetic composition: NVN, NVC, and CVN. N was the nasal consonant /n/. The vowels denoted by V were /i/, /u/, /æ/ and /ɑ/. The consonants denoted by C were /t/ and /s/. Each syllable was spoken in the carrier phrase "say but (nonsense syllable) again." Subjects repeated the randomized speech sample eight times. Prior to reading the speech sample subjects were requested to sustain the sound /m/ in isolation to obtain maximum levels of nasalization for purposes of calibration (Horii and Monroe, 1983).

Data Analysis

Amplitude measurements were made from the N/V ratio signal on the oscillograms. A scale was first determined for each subject, ranging from 0- to 100-percent nasalization. Zero was obtained from the N/V signal value corresponding to the fricative /s/ at the beginning of the carrier phrases on the word "say." For clinical populations exhibiting nasal emission on /s/ an alternative method of determining zero level could be substituted; one could take the zero level from the signal when the accelerometer was disconnected or when the subject was not speaking. One hundred percent nasalization was determined using the sustained /m/. This level also corresponded to the peak nasalization of the /n/s in the speech sample.

Two temporal measurement points were used: (1) the midpoint of the vowel in the nonsense syllables, and (2) a constant time before /n/ onset or after /n/ offset for NVC and CVN items. The beginning and end of the vowel were determined from the microphone signal, using amplitude change and periodicity criteria. The vowel onset and termination points also corresponded with distinctive increases or decreases in the accelerometer signal and the N/V ratio signal amplitude. The midpoint of the vowel was chosen for one of the measurement points, because previous research has indicated that vowel formant frequencies are the most stable at the midpoint of the vowel (Pickett, 1980), and the frequencies of vowel formants are thought to influence the amplitude of nasal accelerometer signals (Stevens et al, 1976). Vowel durations were also measured on the oscillograms. A constant-time measurement point was then chosen for each vowel, corresponding to one-half the mean duration for that vowel. This strategy was chosen, rather than using a single specified time for all the vowels, because vowels are of different intrinsic durations, and we wished to confine measurement points to within the midvowel region.

To determine interjudge reliability of measurement, four repetitions of the speech material were remeasured by another individual familiar with acoustical analysis of speech. A correlation coefficient of 0.98 between the original and replicated measurements was obtained. This value indicates that measurements obtained by this technique are highly reliable between judges.

RESULTS

The data were analyzed statistically using analysis of variance and the Newman Kuels test for multiple comparisons to determine the location of the significant differences.

The first ANOVA was used to test whether the N/V ratio would be greater for the four vowels between two nasal consonants, in comparison to that ratio on a vowel between a nasal consonant and a stop-plosive or fricative. Position of the non-nasal consonant, whether pre- or postvocalic, was not considered as a variable in this first analysis. Only data using midvowel measurement points were included. Figure 1 illustrates these results. The N/V mean signal amplitudes for the four vowels differed significantly from each other ($p < 0.01$), with the exception of means for / æ / and /ɑ/, which did not differ significantly.

Note that the high vowels had a larger N/V signal amplitude than the low vowels (Fig. 1). This finding appears at variance with Moll's (1962) data on vowels adjacent to nasal consonants, since he found velopharyngeal openness greater for / æ / and /ɑ/ than for high vowels. However, our data are consistent with those reported by Stevens et al (1976) for nasal accelerometer signals not divided by a microphone signal, and with those reported by Horii (1986) and Stone and Horii (1986).

Mean data, averaged across vowels, are also shown in Figure 1 for the consonant contexts. The means for the NVN context are significantly greater than for either of the CNV or NVC contexts ($p < 0.01$). Consistent with this finding are results of EMG investigations showing activity of the levator palatini muscle to be greater in vowels adjacent to non-nasal consonants than those adjacent to nasal consonants (Lubker, 1968; Fritzell, 1969; Bell-Berti, 1976). The speech samples in those studies differed from the one used here, however, so a direct comparison cannot be made. There was no significant difference in mean N/V signal amplitudes as a func-

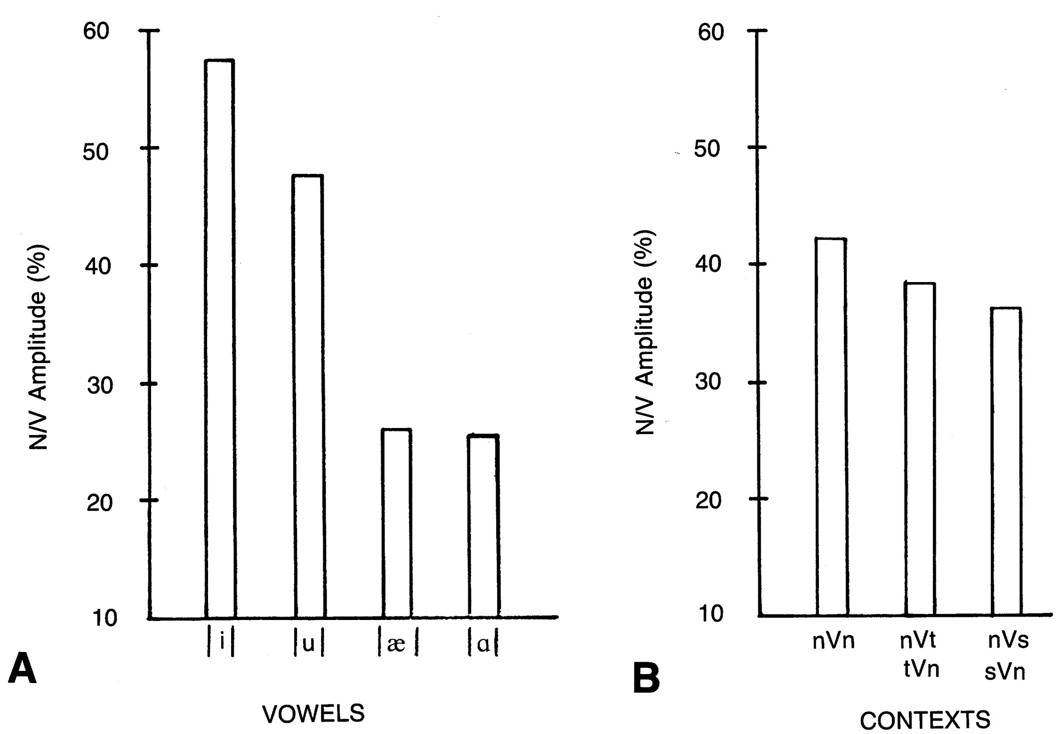

FIGURE 1 *A*, **Mean N/V signal amplitudes for the four vowels** *B*, **Mean N/V signal amplitudes for the consonant contexts.**

tion of manner of production of the non-nasal consonant, whether fricative or stop-plosive.

Mean durations for vowels in the nonsense syllables in this study were 165 msec for /i/, 167 msec for /u/, 217 msec for /ɑ/ and 224 msec for /æ/. The times pre- or postnasal for the constant-time measurement points were one-half the mean duration of the individual vowels.

Further ANOVAs were used to test the effect of pre- or postnasal consonant position and thus whether there was a difference between regressive or progressive assimilation. Data using both measurement points were analyzed separately and the results compared (Table 1). For mid-vowel measurements, only low vowels as a group showed mean N/V ratios higher for prenasal vowels than for post-nasal vowels ($p < 0.01$). Using the constant-time measurement strategy, the data showed a significant difference ($p < 0.001$) for pre- versus postnasal position for all vowel contexts. The N/V ratio amplitudes were larger for the prenasal condition. This last finding is consistent with results of previous research on velar function using radiographic and electromyographic techniques (Moll, 1962; Lubker et al, 1970; Bell-Berti et al, 1979).

DISCUSSION

The results of this study demonstrate that N/V ratios obtained with nasal accelerometer and airborn microphone signals can be a sensitive enough measure of nasal resonance with a vowel to differentiate degree of nasal coarticulation in different phonetic contexts. The findings are generally consistent with results of previous nasal coarticulation studies.

The finding that on the surface appears to contradict the generally accepted relationships between velopharyngeal openness and nasality for different vowels is that N/V ratio amplitudes were greater for high nasalized vowels. Velopharyngeal openness is normally less for high vowels. Since a N/V ratio was used in the present study to adjust for vocal level differences, and high vowels are inherently less intense than low vowels, the N/V ratio would enhance the nasal accelerometer signal magnitude differences between high and low vowels. However, this alone does not explain the effect, since higher accelerometer signal levels have been reported for high vowels even when not adjusted for vocal level.

An acoustical explanation has been offered for why nasal accelerometer signals for high vowels should be greater than for low vowels (Steven et al, 1976). These authors reasoned that the lower the frequency of the first formant, the higher the amplitude of a nasal accelerometer signal. Because high vowels have low frequency first formants and low vowels have high frequency first formants, the accelerometer signal amplitude would be greater for high vowels. Thus, the amplitude of a nasal accelerometer signal (although it reflects degree of *acoustical* coupling), cannot be interpreted as directly reflecting comparative velopharyngeal openness across vowels. The amplitude of accelerometer signals relates rather to perceived nasality in speech contexts not ordinarily nasalized. Therefore, the problem in inferring velar function from a nasal accelerometer signal is similar to the problem of inferring velar function from perceptual judgments of nasality. The accelerometer signal amplitude is a physical measurement, however, and can be more easily quantified.

A promising line for future research would be to obtain simultaneous measurements of accelerometer N/V ratios, vowel formants, and velopharyngeal openness, to see if a formant frequency based weighting factor might be developed that would permit estimates of velar functioning from nasal accelerometer signals. Without such a weighting factor, it is hazardous to infer velar function from N/V ratios when comparing across vowels, and even within a vowel, during any portions of speech involving formant transitions. Thus, the amplitude changes of an accelerometer N/V signal very near a nasal consonant, which might be particularly interesting, are not unambiguously interpretable in respect to physiology at the present time.

The type of measurement made will clearly affect the nature of results when the effects of pre- or postnasalized vowel position are being compared. Sonte and Horii (1986) did not find a significant difference betwen HONC scores for speech material selectively loaded for progressive or regressive assimilation contexts. The HONC index is based on a summation of measurements throughout an utterance and thus cannot compare effects at discrete points in time. The midvowel data from the present study would

TABLE 1 Mean N/V Signal Amplitude for Pre-Versus Postnasal Vowel Contexts for the Two Measurement Strategies

		N/V Signal Amplitude (%)	
Measurement Strategy	*Vowel*	*Prenasal*	*Postnasal*
Midvowel	/i/	55	56
	/u/	46	47
	/æ/	28	20
	/ɑ/	26	21
Constant-time	/i/	59	56
	/u/	54	44
	/æ/	29	20
	/ɑ/	28	21

agree with Stone and Horii's results, except for low vowels. Constant time measurement points yielded results showing a strong position effect, with nasalization for anticipatory coarticulation predominating, as would be predicted by previous physiological work on velar coarticulation. This may imply that a difference in progressive and regressive assimilation, as detected with a nasal accelerometer, is not so much a matter of degree of acoustical coupling near the nasal consonant, but of extent of temporal spread of nasalization.

The utilization of instrumentation to obtain an N/V ratio to be used as feedback in clinical training has found acceptance, because it has been validated by perceptual judgments (Edgerton et al, 1981; Horii and Monroe, 1983; Redenbaugh and Reich, 1985; Reich and Redenbaugh, 1985). The use of temporal detail in the N/V ratio signal to describe or monitor the articulatory aspects of nasalization is promising, subject to the caveats discussed above. More extensive description of expected accelerometer signal characteristics in various phonetic contexts would be useful. Relationships between the temporal fine structure of accelerometric N/V amplitude changes, velar function, and perceived nasal coarticulation deserve further study.

References

Bell-Berti F. An electromyographic study of velopharyngeal function in speech. J Speech Hearing Res 1976; 19:225.

Bell-Berti F, Harris KS, Niimi S. Coarticulatory effects of vowel quality on velar function. Phonetica 1979; 36:187.

Edgerton MT, Sadlove, M, Compton M, Bull G, Blomain V, McDonald E, Bralley R. Nasal vibration system; a non-invasive objective technique to evaluate the speech of patients with palatopharyngeal disorders. Plast Reconstr Surg 1981; 68:153.

Fritzell B. The velopharyngeal muscles in speech: an electromyographic and cineradiographic study. Acta Otolaryng Suppl 1969; 250:1.

Garber SR, Moller KF. The effects of feedback filtering on nasalization in normal and hypernasal speakers. J Speech Hearing Res 1979; 22:321.

Horii Y. An accelerometric approach to nasality measurement; a preliminary report. Cleft Palate J 1980; 17:256.

Horii Y. An accelerometric measure as a physical correlate of perceived hypernasality in speech. J Speech Hearing Res 1983; 26:476.

Horii Y. Effects of rate on nasoaccelerometric amplitudes. Asha 1986; 28:77.

Horii Y, Lang JE. Distributional analyses of an index of nasal coupling (HONC) in simulated hypernasal speech. Cleft Palate J 1981; 18:279.

Horii Y, Monroe N. Auditory feedback of nasalization using a modified accelerometric method. J Speech Hearing Res 1983; 26:472.

Lippmann RP. Detecting nasalization using a low-cost miniature accelerometer. J Speech Hearing Res 1981; 24:314.

Lubker JF. An electromyographic-cinefluorographic investigation of velar function during normal speech production. Cleft Palate J 1968; 5:1.

Lubker JF, Fritzell B, Lindqvist J. Velopharyngeal function: an electromyographic study. Speech Trans Lab Quart Prog Stat Rep 1970; 4:9.

Moll K. Velopharyngeal closure on vowels. J Speech Hearing Res 1962; 5:30.

Pickett JM. The sounds of speech communication: a primer of acoustic phonetics and speech production. University Park Press 1980.

Redenbaugh MA, Reich AR. Correspondence between an accelerometric nasal/voice amplitude ratio and listeners' direct magnitude estimations of hypernasality. J Speech Hearing Res 1985; 28:273.

Reich AR, Redenbaugh MR. Relation between nasal/voice accelerometric values and interval estimates of hypernasality. Cleft Palate J 1985; 22:237.

Stevens KN, Kailkow DN, Willemain TR. A miniature accelerometer for detecting glottal waveforms and nasalization. J Speech Hearing Res 1975; 18:594.

Stevens KN, Nickerson RS, Boothroyd A, Rollins AM. Assessment of nasalization in the speech of deaf children. J Speech Hearing Res 1976; 19:393.

Stone RE, Horii Y. Influences of vowel height and assimilation on HONC dB. Asha 1986; 28:103.