Nasals and nasalization: the relation between segmental and coarticulatory timing

Abstract

Cross-language acoustic and perceptual studies in our lab test the hypothesis that certain aspects of variation in the temporal extent of vowel nasalization are linked to concomitant, inversely related variation in the duration of a flanking nasal consonant. Data from English, Thai, and Ikalanga are reported that support this hypothesis, and possibly related phonological phenomena are considered.

Introduction

The temporal and spatial extent of vowel nasalization varies depending on phonetic context and prosodic structure. For example, in vowels followed by a nasal (N) and then oral (C) consonant, coarticulatory nasalization is more extensive when C is voiceless than when it is voiced (e.g., Malécot, 1960) or when C is a fricative (Ohala & Busà, 1995; Busà, in press). Prosodic influences on vowel nasalization include greater nasalization in tautosyllabic than heterosyllabic VN sequences (Cohn, 1990; Solé, 1995) and in stressed than unstressed syllables (Vaissière, 1988; Krakow, 1993). The height and duration of the vowel itself also matter, with lower (Bell-Berti, 1993) and longer (Whalen & Beddor, 1989) vowels tending to be more heavily nasalized in nasal contexts.

In recent years, research in our lab concerning contextual variation in vowel nasalization has been unified by the hypothesis that some aspects of variable vowel nasalization are due to variation in the temporal alignment of the nasal and oral gestures for N. Under this hypothesis, more extensive vowel nasalization in productions of, for example, /Vns/ compared to /Vnz/ or /Vnt/ is due not to an increase in nasalization per se but rather to earlier onset of a roughly constant-sized nasal gesture relative to tongue-tip raising for /n/ in /Vns/. If this hypothesis is correct, then production measures should show a trade-off in the relative durations of vowel nasalization and N: the duration of N should be inversely related to the extent of its coarticulatory influence on a flanking vowel.[1]

This paper provides an overview of our experiments with American English, Botswanan Ikalanga, and Thai as they pertain to the temporal interplay between vowel nasalization and N duration. The languages were chosen in large part because they differ in the segmental contexts in which VN sequences are phonotactically permissible. The primary focus is on production data, which will be shown to be generally consistent with the temporal alignment hypothesis, although some languages show little effect of context on nasals or nasalization. Reported in less detail is cross-language research on perception of the interplay between Ṽ and N testing whether, given that vowel nasalization and a flanking nasal are in a temporal trading relation in production, listeners hear these cues for nasality as perceptually equivalent.

Temporal relations between Ṽ and N in production

Measurements

Four to six speakers of each language were recorded reading randomized word lists targeting VN sequences, with the purpose of the lists being masked by fillers. Acoustic measures included vowel duration, duration of vowel nasalization, nasal consonant duration, and

[1] The temporal alignment hypothesis, and the expected inverse relation between vowel nasalization and N duration, is predicted to hold only for (certain) effects of segmental context. That the velum gesture is temporally and spatially larger in some prosodic positions than others is well established (Krakow, 1993, among others).

duration of flanking oral consonants. Acoustic onset of vowel nasalization was determined by inspecting FFT spectra in 10 ms increments across the vowel; onset was identified as the first spectrum with an identifiable low-frequency nasal formant and/or a broadening of F1 bandwidth and lowering of F1 amplitude. For some speakers nasal and oral airflow data were collected in separate recordings, although only acoustic data are reported here; the focus on acoustic measures follows from our dual interest in the production and perception of timing relations.

Effects of obstruent voicing on temporal alignment

In some languages, nasal consonants are substantially shorter before voiceless than before voiced consonants; of interest here is whether shorter N co-occurs with greater vowel nasalization. The hypothesized relation for VNC sequences is shown schematically in Figure 1, where earlier initiation of the velum gesture in a voiceless context (lower panel) would, relative to a voiced context (upper), yield (a) temporally more extensive vowel nasalization, (b) a shorter acoustic nasal murmur, and (c) longer post-nasal oral constriction.

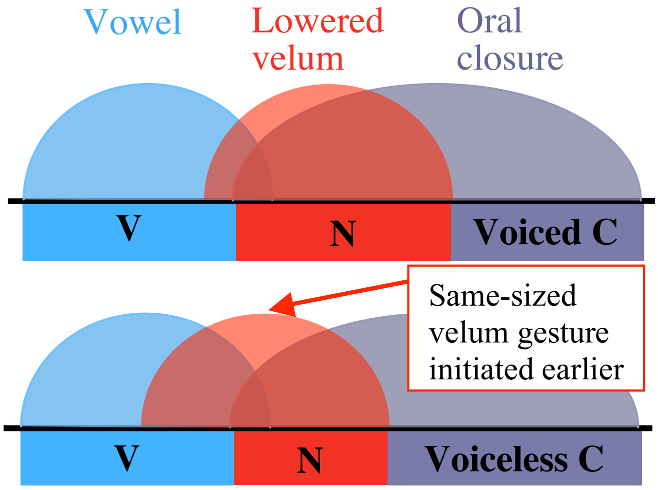

Figure 1: Schematic representation of consequences for vowel nasalization, nasal murmur, and post-nasal oral constriction if velum gesture is initiated earlier in voiceless than voiced contexts.

The temporal measures described above were applied to VNC sequences produced by speakers of English and Ikalanga. For English, 5 speakers read word lists containing /ɛnC/ (e.g., *spend*, *spent*, *dens*, *dense*). English shows extreme N shortening in pre-voiceless position (e.g., Raphael et al., 1975); we investigated whether N shortening results in greater vowel nasalization and a longer post-nasal C. Figure 2 plots the relation between N duration and duration of vowel nasalization for approximately 50 tokens of /ɛnC/ from each of the 5 speakers. Trendlines calculated for each speaker's productions are also given; R^2 statistics ranged from .27 to .45. Although not all of the variation in vowel nasalization is accounted for by N duration—the total nasalization ($\tilde{V}$ plus N) being greater in voiced than in voiceless contexts—the predicted inverse relation between $\tilde{V}$ and N durations holds for all speakers (see also Malécot, 1960 and Cohn, 1990). Moreover, this relation holds not only across voicing contexts, but also for finer-grained temporal differences within the voiceless context. That is, a scatterplot (not given here) of only $VNC_{[voiceless]}$ tokens shows an inverse relation similar to that in Figure 2, albeit with somewhat less steep trendlines ($R^2 \approx 0.2$ for $VNC_{[voiceless]}$). That the co-variation is due in part to earlier onset and earlier offset of the velum gesture relative to the oral cavity configuration is supported as well by C durations (also not plotted here to conserve space), which are inversely related to N duration in the voiceless context: the shorter the N in VNC$_{[voiceless]}$, the longer the post-nasal oral constriction.[2]

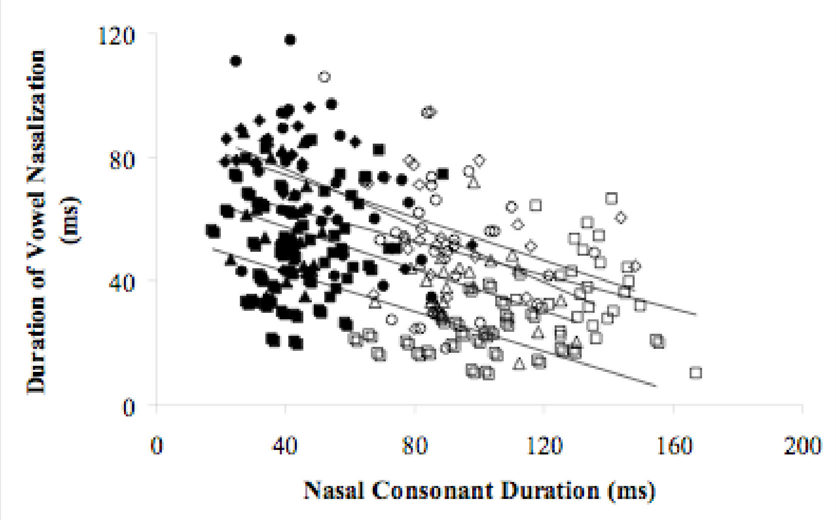

Figure 2. Scatterplot showing inverse relation between nasal consonant and vowel nasalization duration for /ɛnC/ tokens from 5 American English speakers. Symbol shape denotes tokens from a given speaker. Closed symbols = VNC$_{[voiceless]}$ and open = VNC$_{[voiced]}$. A trendline is provided for each speaker.

A different timing pattern emerges for VNC sequences in Ikalanga, which show no influence of voicing on nasalization. In Ikalanga, as in other Bantu languages, NC sequences are traditionally analyzed as prenasalized ^{N}C (Mathangwane, 1999; but see Downing 2005 for a cluster analysis). Most VNCVs in Iklanaga are voiced, although NC$_{[voiceless]}$ occur in borrowed words, including some common words. In the experiment, 6 Botswanan Ikalanga speakers read word lists that included voiced VNCV (e.g., [da^{n}da]), voiceless VNCV (e.g., [ke^{n}ta]), and voiced NCV (e.g., [^{n}dulo]; initial voiceless NC do not occur). The resulting temporal measures in Figure 3 give no evidence of N shortening in pre-voiceless contexts and, perhaps consequently, show no increase in vowel nasalization in NC$_{[voiceless]}$ relative to NC$_{[voiced]}$ contexts. While there is slight evidence of an inverse relation between Ṽ and N durations in Ikalanga, contrary to English it is the voiceless context that has longer N and less extensive vowel nasalization. Total nasalization (Ṽ plus N durations) is the same for the voiced and voiceless VNCV contexts (122 ms); the bottom panel shows that the duration of initial N (in NCV), where there is no preceding vowel, is the same (121 ms) as Ṽ plus N duration for VNCV sequences.

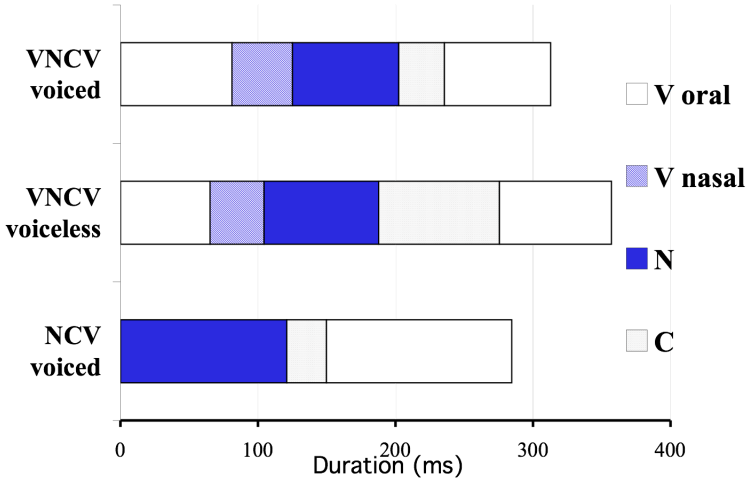

Figure 3. Durations of oral and nasalized portions of vowels, and of nasal (N) and oral (C) consonants, in VNCV and NCV sequences, averaged across productions of 6 Ikalanga speakers. Duration of acoustic nasality (Ṽ+N in VNCV) is constant across the sequence types.

[2] Abby Cohn and Sarah Hawkins provided valuable insights into the notion of temporal variation in VNC sequences as due to "sliding" of the nasal gesture. That this notion better accounts for variation within voiceless than within voiced contexts merits further study.

In summary, while only English VNCV sequences showed clear evidence of the predicted inverse relation between N duration and the temporal extent of vowel nasalization, the Ikalanga data nonetheless suggest a relatively constant-sized nasal gesture across VNCV (and NCV) sequences. It is noteworthy that the existing literature provides additional evidence of $\tilde{V}$ – N co-variation in VNC sequences triggered not by voicing, but by frication. Busà (in press) found that, in some Italian dialects, N is shorter and vowel nasalization temporally more extensive in $VNC_{[fricative]}$ than $VNC_{[stop]}$ sequences, and Hattori et al. (1958) reported a similar pattern in Japanese.

Effects of vowel length on the relation between $\tilde{V}$ and N

A second context effect investigated in our studies, in this case for Thai and English, is the influence of vowel length on the temporal relation between vowel nasalization and flanking nasals. Thai has contrastive vowel length and nasal coda duration in Thai is inversely related to vowel duration (Roengpitya, 2001). Onsuwan (2005), in our lab, investigated whether N duration in Thai was also inversely related to the temporal extent of coarticulatory vowel nasalization. Using the acoustic measures described above, Onsuwan analyzed an extensive set of CV(:)N (e.g., /bén/, /be:n/, /bam/, /bà:m/) words and nonsense items produced by 4 Thai speakers. In a separate study, we investigated essentially the same question for English tense and (shorter) lax vowels for productions of CVN words (e.g., *seen*, *sin*, *pain*, *pen*) by 6 American English speakers (Sefton and Beddor, 2005). Figure 4 presents the overall results for Thai and English: the VN contexts with the longer nasal codas (short VN in Thai and lax VN in English) co-occur with significantly less nasalized vowels. Although the total duration of acoustic nasalization ($\tilde{V}$ plus N) is not precisely constant across long (tense) and short (lax) vowel contexts, VN sequences in both languages exhibit the predicted trade-off in the relative durations of vowel nasalization and N.

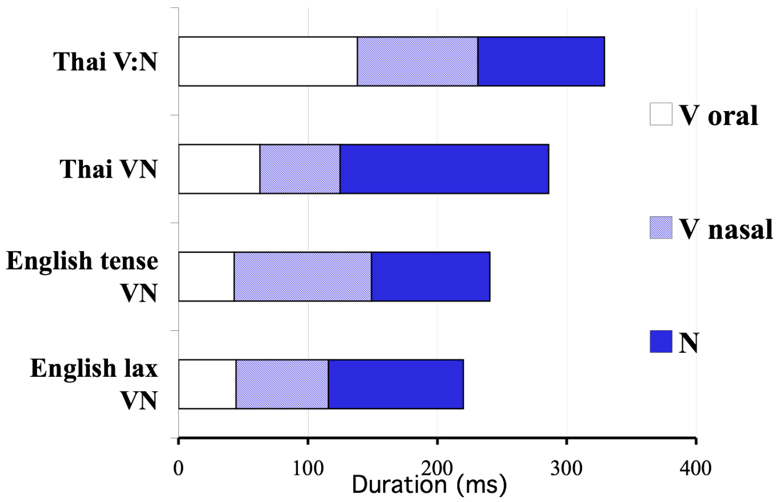

Figure 4: Durations of oral and nasalized portions of vowels and of nasal consonants (N) in CVN sequences, averaged across productions of 4 speakers of Thai (upper bars) and 6 speakers of English (lower bars). The Thai data are from Onsuwan (2005).

Perceived equivalence of nasals and nasalization

Thus our production studies, together with the existing literature, support the hypothesis that some aspects of contextual variation in vowel nasalization are linked to inversely related variation in the duration of a following nasal consonant. These findings have led us to hypothesize that, in the face of the articulatory interplay between $\tilde{V}$ and N, listeners will treat nasality on vowels and consonants as perceptually equivalent. To test this hypothesis, we edited natural tokens of Ikalanga /gaba/ and /gamba/, which sound like perfectly acceptable possible but non-occurring words in English, creating (a) a 9-step N continuum from 0 to 70 ms of [m] murmur and (b) two degrees of vowel nasalization, slight nasalization (where the final 20% of /a/ was nasalized) and heavier (final 50% nasalized). Using a trading relations paradigm (Fitch et al., 1980), three types of discrimination pairs were created. All stimulus pairings had a constant-sized (34 ms) difference in N duration. *Nasal-only* pair members, N_S-N_L (where $_S$ = shorter and $_L$ = longer N), differed only in N duration (i.e., [gãm$_S$ba]-[gãm$_L$ba]. *Different nasality pairings*, $\tilde{V}_S N_S$- $\tilde{V}_L N_L$, differed in $\tilde{V}$ and N, with one pair member having slight vowel nasalization and shorter N, and the other heavier nasalization and longer N ([gã$_S$m$_S$ba]-[gã$_L$m$_L$ba]). *Similar nasality* stimuli, $\tilde{V}_S N_L$- $\tilde{V}_L N_S$, also differed in $\tilde{V}$ and N but with the opposite pairing of these manipulations so that total nasalization across the VN sequence was more nearly constant ([gã$_L$m$_S$ba]-[gã$_S$m$_L$ba]). (See Beddor et al., in press, for a more detailed presentation of the perceptual paradigm, albeit for a different set of stimuli.)

If listeners hear vowel and consonant nasality as perceptually equivalent they should poorly discriminate the *similar nasality* trials, and be more accurate in discriminating *different nasality* trials, even though the two trial types have the same-sized acoustic differences between pair members (i.e., in both, vowel nasality differed by 30% and N duration by 34 ms). The predicted outcome should hold especially for speakers of a language (e.g., English) in which $\tilde{V}$ and N systematically co-vary in production, although we speculated that, all listeners, even if their language has little $\tilde{V}$-N co-variation (e.g., Ikalanga), might show some evidence of perceived equivalence. Figure 5 gives the pooled results for 23 American English and 24 Botswanan Ikalanga speakers, summing across the multiple stimulus pairings (along the N continuum) of each of the three pair types. Listeners from both language backgrounds show the predicted pattern of perceived equivalence, being least accurate on trials in which total nasality across VN was roughly constant ($\tilde{V}_L N_S$-$\tilde{V}_S N_L$)—less accurate even than for N-only (N_S-N_L) pairings, which have smaller acoustic differences between stimuli. Not surprisingly, American English listeners, exposed to systematic $\tilde{V}$-N co-variation in their language, had greater difficulty than Ikalanga listeners discriminating stimuli varying only in N (N_S-N_L) or with similar nasality ($\tilde{V}_L N_S$- $\tilde{V}_S N_L$).

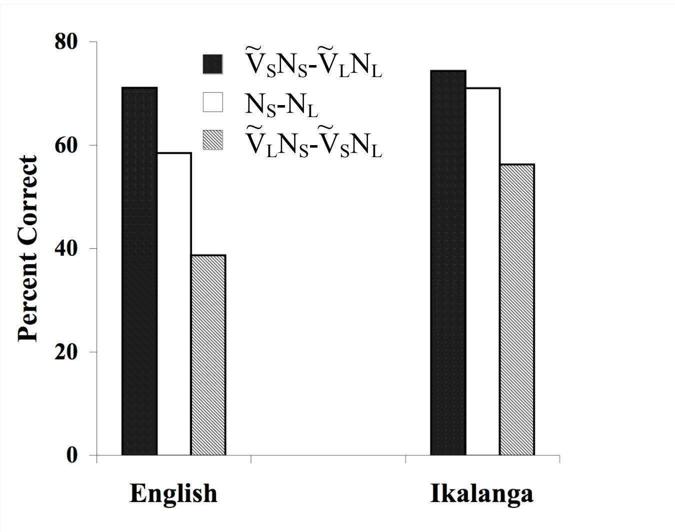

Figure 5. Pooled responses of 23 American English and 24 Ikalanga listeners to three types of discrimination trials (see text).

Discussion and phonological implications

Although nasalized vowels and nasal consonants show substantial temporal variation across contexts, our production data indicate that the acoustic manifestations of nasal gestures in VN sequences have some temporal stability: contexts that trigger shorter nasal consonants have concomitantly longer anticipatory vowel nasalization. The perceptual consequence of (relative) temporal stability of nasalization across the VN sequence is that listeners are more sensitive to acoustic variation in total nasalization ($\tilde{V}$+N) than to extent of nasalization on $\tilde{V}$

or N, and experimentally respond to vocalic and consonantal nasality as though they were perceptually equivalent.

Several widely attested phonological patterns involving nasal codas are in keeping with these findings. Phonologically, nasal codas are robust in that languages that allow syllable codas nearly always allow coda nasals, and some languages allow only nasals as codas (e.g., Davis, 2002). In addition, languages that have lost (some or all) coda consonants and for which a detailed chronology of that loss can be determined (e.g., Chinese and Romance languages; Chen and Wang, 1975) show diachronically and synchronically that the loss of N follows a slower trajectory than that of oral stops. While converging factors are likely at play in these phenomena, one factor may be the temporal stability of coda nasality across VN. Perceptual equivalence of nasality on Ṽ and N should further contribute to N coda robustness, as even an extremely short nasal murmur (with accompanying vowel nasalization) would perceptually "count" as N. At the same time, co-variation between Ṽ and N in production, together with their tight perceptual link, should mean that listeners—especially learners—might arrive at /Ṽ/ as the representation for speakers' intended and even actual /VN/, an expectation that is borne out by the historical evolution of Ṽ from VN in many languages. Thus acoustic, perceptual, and phonological findings converge in showing the systematic interplay between segmental and coarticulatory timing for nasals.

Acknowledgments

This research was supported by NSF Grant BCS-0118684. A subset of the results of the English /ɛnC/ data is reported by Beddor et al. (in press). The assistance of Anthony Brasher and Chutamanee Onsuwan, and the helpful comments of the University of Michigan phonetics-phonology research group, are gratefully acknowledged.

References

Beddor, P. S., Brasher, A., and Narayan, C. (In press). 'Applying perceptual methods to the study of phonetic variation and sound change', in M. J. Solé, P. S. Beddor, and M. Ohala (eds.), *Experimental Approaches to Phonology*. Oxford.

Bell-Berti, F. (1993). 'Understanding velic motor control: studies of segmental context,' in M. K. Huffman and R. A. Krakow (eds.), *Nasals, Nasalization, and the Velum.* New York, NY: Academic Press, 63-85.

Busà, M. G. (In press). 'Coarticulatory nasalization and phonological developments: data from Italian and English nasal-fricative sequences', in M. J. Solé, P. S. Beddor, and M. Ohala (eds.), *Experimental Approaches to Phonology*. Oxford.

Chen, M. Y. and Wang, W. S-Y. (1975). 'Sound change: actuation and implementation', *Language* 51, 255-281.

Cohn, A. (1990). 'Phonetic and phonological rules of nasalization,' *UCLA Working Papers in Phonetics* 76: 1-224.

Davis, S. (2002). 'Syllable structure for an artificial language based on universal principles', *Journal of Universal Language* 3, 3-15.

Downing, L. J. (2005). 'On the ambiguous segmental status of nasals in homorganic NC sequences', in M. Van Oostendorp and J. M. van de Weijer (eds.), *The Internal Organization of Phonological Segments.* Berlin: Mouton de Gruyter.

Fitch, H. L., Halwes, T., Erickson, D. M., and Liberman, A. M. (1980). 'Perceptual equivalence of two acoustic cues for stop-consonant manner', *Perception & Psychophysics* 27: 343-350.

Hattori, S., Yamamoto, K., and Fujimura, O. (1958). 'Nasalization of vowels in relation to nasals', *Journal of the Acoustical Society of America* 30: 267-274.

Krakow, R. A. (1993) 'Nonsegmental influences in velum movement patterns: syllables, sentences, stress, and speaking rate', in M. K. Huffman and R. A. Krakow (eds.), *Nasals, Nasalization, and the Velum.* New York, NY: Academic Press, 87-113.

Malécot, A. (1960). 'Vowel nasality as a distinctive feature in American English', *Language* 36: 222-229.

Mathangwane, J. T. (1999). *Ikalanga Phonetics and Phonology: A Synchronic and Diachronic Study.* Stanford, CA: CSLI Publications.

Ohala, J. J. and Busà, M. G. (1995). 'Nasal loss before voiceless fricatives: a perceptually-based sound change', *Rivista di Linguistica* 7: 125-144.

Onsuwan, C. (2005). *Temporal relations between consonants and vowels in Thai syllables*. University of Michigan doctoral dissertation.

Roengpitya, R. (2001). *A study of vowels, diphthongs, and tones in Thai*. University of California, Berkeley, doctoral dissertation.

Sefton, S. and Beddor, P. S. (2005). 'Nasals and nasalization: the interplay between segmental duration and coarticulation.' Paper presented at the 149th Meeting of the Acoustical Society of America, Vancouver.

Solé, M.-J. (1995). 'Spatio-temporal patterns of velopharyngeal action in phonetic and phonological nasalization', *Language and Speech* 38, 1-23.

Vaissière, J. (1988). 'Prediction of articulatory movement of the velum from phonetic input,' *Phonetica* 45, 122-139.

Whalen, D. H. and Beddor, P. S. (1989). 'Connections between nasality and vowel duration and height: elucidation of the Eastern Algonquian intrusive nasal,' *Language* 65: 457-486.