

Perception of Prenasalized Stops

Patrice Speeter Beddor and Chutamane Onsuwan

University of Michigan, Ann Arbor, MI USA 48109

E-mail: beddor@umich.edu, consuwan@umich.edu

ABSTRACT

Prenasalized stops are estimated to occur in 10-15% of the world's languages, yet there has been little investigation of their perceptual characteristics. A series of acoustic and perceptual experiments was conducted on the prenasalized stops of the Bantu language Ikalanga. The main acoustic cues differentiating prenasalized stops (^NC) from plain nasals (N) were a short post-nasal oral closure for ^NC and heavy coarticulatory vowel nasalization following N. Independent manipulation of these cues showed that Ikalanga listeners used both segmental and coarticulatory timing in identifying ^NC vs. N, although carryover coarticulatory nasalization was both a sufficient and necessary cue for the distinction. Comparable weighting of anticipatory nasalization was not found for tests of the ^NC -C distinction, although this different perceptual outcome is in keeping with Ikalanga acoustic patterns.

1. INTRODUCTION

Our investigation of prenasalized stops is part of a larger study of the relation between segmental and coarticulatory timing in production and perception. Prenasalized stops differ from nasals in the timing of velum raising relative to the release of the oral closure: the velum remains lowered throughout the oral closure in plain nasals (N), but raises prior to release of the oral closure in prenasalized stops (^NC). Of special interest here, given our focus on relative timing, is that the "C" component, at least for voiced ^NC s, tends to be quite short. For example, Maddieson [1] reported that the oral stop portion of ^NC s in Fijian was very brief, and not reliably measurable. Measures of voiced ^NC s in several Bantu languages indicate that the N portion of ^NC is about 4 times longer than the C portion [2, 3], and the same holds for our measures of Sinhala ^NC s.

The small difference in velic timing between voiced ^NC and N closures is often accompanied by other temporal differences. Many Bantu languages exhibit vowel lengthening before ^NC s but not Ns [3]; in some Bantu languages, the nasal murmur of ^NC is also longer than that of N [2]. ^NC and N also differ in coarticulatory timing of velum raising beyond the oral closure, and this difference can be quite prominent. Our preliminary work on Sinhala indicates heavy carryover nasalization after N but not ^NC . Connell [4] has made similar observations for Mambila and, to a lesser extent, Ibibio, although Maddieson [5] noted that carryover nasalization after N is not particularly prominent in, for example, Sukuma or Fijian.

This study explores the relative contributions of segmental and coarticulatory timing differences to the perception of the contrast between ^NC and N. Our investigation focuses on Ikalanga, a Bantu language of Botswana and Zimbabwe. The Ikalanga contrast between ^NC and N is of particular interest because the only systematic difference in segmental timing is that, in voiced ^NC s, the velum raises a short 10-25 ms prior to the release of the oral closure. The difference between ^NC and N closures is accompanied by a difference in post-closure coarticulatory nasalization, with the lowered velum position for N extending throughout much of the following vowel.

2. ACOUSTICS OF IKALANGA ^NC s

Ikalanga has a three-way contrast in initial and medial position among ^NC , N, and voiced stops C (e.g., [gá^mbá] 'frozen', [gamá] 'you catch', and [gabá] 'bucket'). Setting aside possible questions related to the analysis of ^NC s, Ikalanga has exclusively open syllable structure [6]. As a preliminary step to our perceptual study, we recorded 6 Botswanan Ikalanga speakers reading multiple randomizations of a 100-word reading list containing (V) ^NC V, (V)NV, and (V)CV, sequences (where ^NC = [^mb ⁿd, ^ŋg], N = [m n ŋ]), and C = [b d g]). Fig. 1 gives the results, averaged across speakers, of several temporal measures taken for the disyllabic sequences. The temporal extent of coarticulatory vowel nasalization was measured by inspecting FFT spectra in 10 ms increments throughout the course of the vowel. Offset (or onset) of carryover (anticipatory) nasalization was identified as the last (first) display with a clear low-frequency nasal formant and/or a broadening of F1 bandwidth and lowering of F1 amplitude.

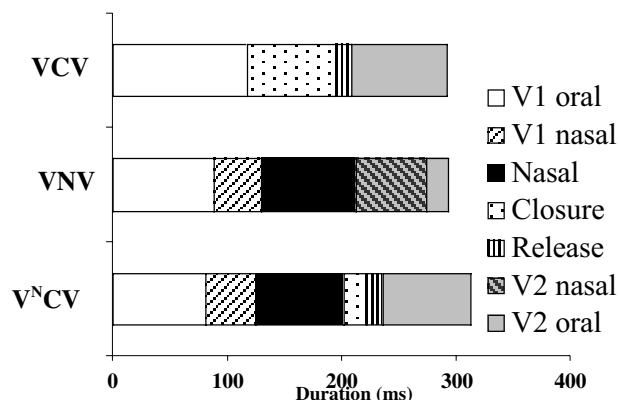


Figure 1. Segmental and coarticulatory timing in Ikalanga VCV, VNV, and V N CV, averaged across 6 speakers.

Ikalanga has penultimate stress, giving rise to longer initial than final vowels in the disyllables. However, as shown by the bottom two bars of Fig. 1, vowel duration (i.e., V1 oral + V1 nasal or V2 oral + V2 nasal) does not change depending on whether the medial consonant is ^NC or N. Nasal murmur durations of N and ^NC are also the same, falling within 5 ms of each other. As noted in the Introduction, the main difference in segmental timing between ^NC and N is that ^NC has a post-nasal oral occlusion plus release burst. This occlusion averages one-fourth the duration of the nasal murmur (for ^NC, ^N = 77 ms, C = 19 ms), and is also one-fourth the duration of the closure for voiced oral stops (C, top bar). Coarticulatorily, Ikalanga ^NC and N trigger the same pattern of moderate anticipatory nasalization on a preceding vowel (V1 nasal). N, but of course not ^NC, also triggers temporally extensive carryover nasalization extending through 75% or more of a following vowel (V2 nasal). In addition, Ikalanga VNV sequences have *spatially* extensive carryover nasalization. Although not indicated in Fig. 1, the A1-AN difference (amplitude of highest-amplitude harmonic of F1 minus that of FN [7]) was smaller—indicating greater low-frequency spectral flatness—for vowels with carryover than those with anticipatory nasalization.

3. PERCEPTION OF THE ^NC – N CONTRAST

Guided by the acoustic findings for Ikalanga V^NCV and VNV sequences, we designed an identification experiment to investigate the relative contributions of the oral occlusion + burst versus carryover vowel nasalization to the perception of the ^NC vs. N distinction. The perceptual stimuli were created by applying waveform editing techniques to tokens of low-toned /ga^mba/, /gama/, and /gaba/ produced by a female speaker of Ikalanga. The original tokens were selected on the basis of goodness of match for pitch, amplitude, and formant frequencies. These tokens were edited to create a 16-step /ga^mba-gama/ (phonetically, [ga^mba-gamã]) series. Each edited stimulus in the series consisted of three portions. The initial /gam/ (/ga^m/), excised from a token of /ga^mba/ with a 90-ms nasal murmur, was constant across all stimuli. The medial /b/ portion varied across stimuli and consisted of a 0, 9, 18, or 27 ms splice excised from a 27-ms post-nasal oral occlusion + burst from /ga^mba/. (The 9 and 18 ms splices included the short burst.) The final /a/ portion also varied across stimuli and was created by replacing, in 29-ms increments, portions of the oral final vowel of /gaba/ with same-sized splices from the heavily nasalized final vowel of /gama/, yielding final vowels that were 0, 36, 68, or 100% nasalized. (The final vowel of /gaba/, rather than /ga^mba/, was chosen because a more nearly identical match with /gama/ was found in terms of formant frequencies.) In concatenating the three portions, minor amplitude adjustments were sometimes made at the precise location of the splice to avoid abrupt spectral discontinuities in the signal. The splicing process introduced no audible discontinuities in the signal and stimulus naturalness was verified with a native speaker of Ikalanga.

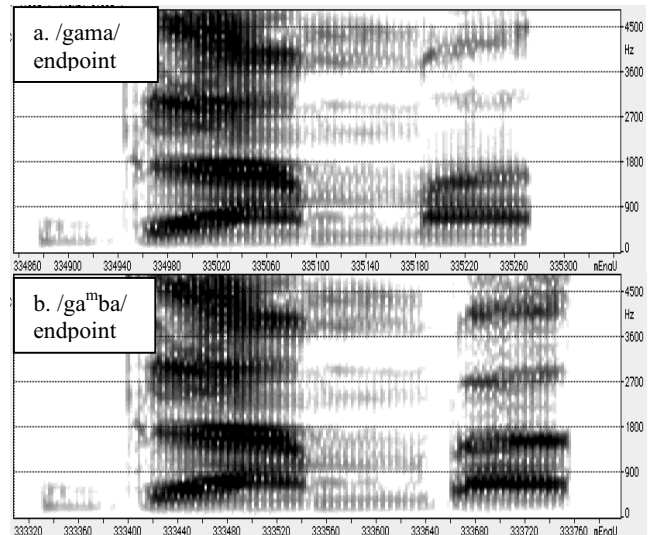


Figure 2. Spectrograms of the /gama/ (a) and /ga^mba/ (b) endpoints of the 16-step /ga^mba-gama/ series.

Covarying the 4-step "segmental" continuum (duration of [b] portion) and the 4-step "coarticulatory" continuum ([a-ã]) yielded the 16 /ga^mba-gama/ stimuli. In Fig. 2, the top spectrogram shows the endpoint /gama/ stimulus with no post-nasal oral closure but with a fully nasalized final vowel. The bottom spectrogram shows the endpoint /ga^mba/ stimulus, with a 27-ms post-nasal oral closure + burst and an oral final vowel. The segmental and coarticulatory timing for the endpoint stimuli was consistent with natural speech productions of this speaker.

The listeners were 29 native speakers of Ikalanga who were students at the University of Botswana. (Thirty listeners participated in the testing, but one listener had difficulty with the task, and those data are excluded from the results presented below.) Ten randomizations of the stimuli were presented in a self-paced identification task. Listeners identified each item as either /ga^mba/ or /gama/ and recorded their response by clicking on the appropriate button ("gamba" or "gama") on the computer screen. After recording each response, listeners selected "next trial" when they were ready to hear the next stimulus. The test was presented over headphones to groups of three listeners (one listener per laptop) in a quiet room at the University of Botswana.

The identification functions in Fig. 3 give the percent ^mb/ responses to the /ga^mba-gama/ stimuli, according to the segmental continuum (x axis), pooled across listeners. By way of orientation, we note that the /gama/ endpoint in Fig. 2a corresponds to the right-most open circle at the bottom of Fig. 3; the /ga^mba/ endpoint in Fig. 2b corresponds to the left-most diamond at the top of Fig. 3.

Clearly, identification of the distinction between /m/ and ^mb/ depended on both segmental and coarticulatory timing. This overall outcome is in keeping with the general finding in perception studies that listeners are sensitive to the rich information provided by the acoustic

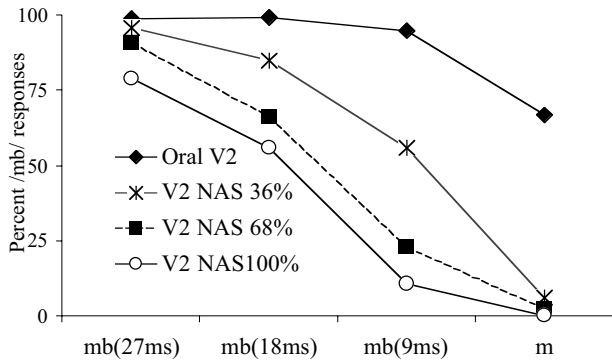


Figure 3. Pooled percent /^mb/ responses to the 16-step /ga^mba/-/gaba/ series, according to the segmental continuum.

signal about what the speaker has said. Importantly, however, the segmental variation alone, without accompanying carryover vowel nasalization, was not sufficient to cue to the /m/-^mb/ distinction: no stimulus in the "oral V₂" series (i.e., the diamonds in Fig. 3) was reliably identified as /m/.

In contrast, at least for these stimuli, the variation in coarticulatory vowel nasalization was both *necessary and sufficient* for listeners to differentiate /m/ and /^mb/. That some carryover nasalization was necessary for listeners to hear /m/ follows from the insufficiency of the segmental cue. That carryover nasalization alone, with no post-nasal oral occlusion plus burst, was sufficient to cue the /m/-/^mb/ contrast is shown by the four right-most stimuli in Fig. 3 (the four "m" stimuli). Although, not surprisingly, the nasal-only "m" series was not as convincing as the series with 9 ms post-nasal oral closure, the [a-ã] variation in the "m" series provided listeners with sufficient information to reliably differentiate /m/ from /^mb/.

We expect that the crucial role of coarticulatory timing for perception of the contrast between ^NC and N is due to the temporal and spectral magnitude of the coarticulatory effects in production, together with the shortness of the post-nasal oral closure (which is on the order of 10 ms for some speakers). By way of comparison, we conducted a follow-up experiment which tested the contribution of segmental and coarticulatory timing to the perception of contrast between ^NC and voiced oral C in Ikalanga.

4. PERCEPTION OF THE ^NC – C CONTRAST

Our perceptual investigation was again guided by the acoustic findings, this time for Ikalanga V^NCV and VCV sequences. As shown in Fig. 1, ^NC and C exhibit large segmental differences in terms of velum position during the oral occlusion. On the other hand, the coarticulatory difference between ^NC and C is less extensive than for the ^NC-N contrast, with anticipatory nasalization extending, on average, through one-third of the vowel preceding ^NC. A1-AN measures were also consistent with less prominent anticipatory than carryover nasalization.

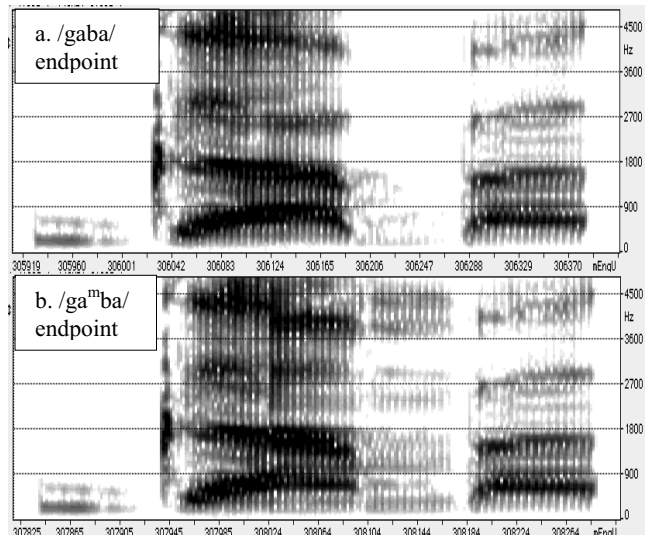


Figure 4. Spectrograms of the /gaba/ (a) and /ga^mba/ (b) endpoints of the 36-step /gaba-ga^mba/ series.

We created a /gaba-ga^mba/ series which again covaried segmental and coarticulatory timing, although in this case the coarticulatory variation was in the initial rather than the final vowel. Original tokens of low-toned /gama/ and /gaba/, produced by the same female Ikalanga speaker, were edited to create stimuli having three concatenated portions. The initial /ga/ was a stimulus from a 4-step "coarticulatory" continuum created by incrementally replacing, from right to left, portions of the oral initial vowel of /gaba/ with same-sized splices from the nasalized initial vowel of /gama/, yielding initial vowels that were 0, 20%, 36%, and 52% nasalized. The medial /b/ or /^mb/ was a stimulus from a 9-step "segmental" continuum created by replacing, from left to right in 6-12 ms increments, portions of oral pulsing for /b/ from /gaba/ with same-sized splices from the nasal murmur of /gama/. (Nasal murmur increments were 6 ms for the first 4 steps of the continuum and 12 ms thereafter.) The final /a/, excised from /gaba/, remained constant across all stimuli. Fig. 4 gives spectrograms of the /gaba/ and /ga^mba/ endpoints of the resulting 36-step series. The /ga^mba/ endpoint had 52% anticipatory nasalization of the initial /a/ followed by a 70-ms nasal murmur and 20 ms of oral pulsing + burst.

The identification test consisted of 10 randomizations of the 36 stimuli. Procedures were as in the previous experiment, except that the same 29 native speakers of Ikalanga identified each token as /gaba/ or /ga^mba/.

Fig. 5 gives the pooled percent /^mb/ responses to the /gaba-ga^mba/ stimuli. Both segmental and coarticulatory cues affected listeners' identifications, but the segmental variation alone (the diamonds) was sufficient for systematic categorization of the stimuli as /b/ or /^mb/. Indeed, listeners' sharpest crossover from /b/ to /^mb/ (i.e., with the fewest ambiguous stimuli) was in their responses to the oral vowel continuum. Thus, unlike the contribution of carryover nasalization to perception of the /^mb/-/m/ distinction, anticipatory nasalization was not

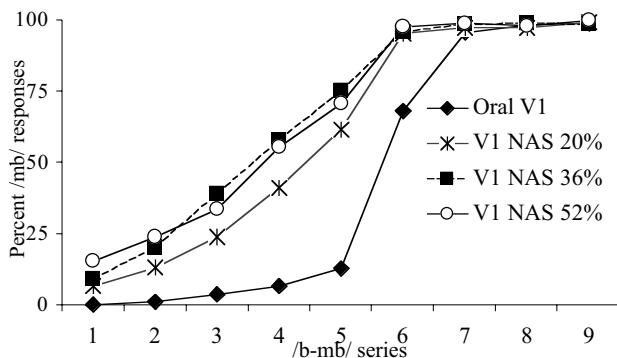


Figure 5. Pooled percent ^mb/ responses to the 36-step /gaba-/ga^mba/ series. Stimulus 1: 0 ms ^m/ and 90 ms /b/ portion; stimulus 9: 70 ms ^m/ and a 20 ms /b/ portion.

necessary for perception of the /b-/^mb/ distinction. Of course, this overall outcome is not surprising in view of the large segmental variation in the /gaba-/ga^mba/ series, which is in turn consistent with Ikalanga acoustic patterns.

5. SUMMARY AND CONCLUSIONS

^NCs in the world's languages differ from plain Ns in that velum raising occurs prior to release of the oral closure in ^NC, but may occur well after the oral closure for N. In Ikalanga, velum raising after N is considerably delayed with, on average, 75% of the following vowel being strongly nasalized. Perceptual testing that independently manipulated nasalization during the oral closure and during the post-closure vowel showed not only that Ikalanga listeners' judgments of medial consonants as ^NC or N were influenced by carryover vowel nasalization, but that (at least for these stimuli) identification as N required some carryover nasalization. On the other hand, when the vowel was oral, listeners reported hearing ^NC even when the preceding closure was completely nasal. In these respects, the perceptual weight of carryover nasalization for the Ikalanga ^NC-N distinction goes beyond the context effects typically found for coarticulatory information. The prominent role of coarticulatory nasalization in the perception and production of the contrast between N and ^NC in Ikalanga bears some resemblance to historical changes involving ^NCs and nasal + stop clusters. For example, in Acehnese, proto ^NCs evolved into orally released nasals while plain Ns in Acehnese trigger heavy carryover nasalization [3, 5]; see [8] for similar cases.

Whether similar importance of coarticulatory timing for the perception of ^NC-N contrasts holds for other languages with prenasalized stops remains to be determined. We expect that perceptual weighting will largely mirror production in a given language. In the current investigation, we take as support for this position the different behavior, in both perception and production, of the ^NC-C contrast. We are also now investigating perception of ^NC and N in Sinhala, a language with extensive carryover coarticulation after N and extensive anticipatory coarticulation before ^NC (as well as N).

While we do not yet know whether similar perceptual findings will hold across languages with different timing patterns for ^NCs, there is also no reason to believe that the patterns reported here are unique to languages with phonological ^NCs. Browman and Goldstein [10], for example, reported similar articulatory patterns for Kichaga ^NCs and English (phonological) nasal + stop clusters. It is therefore not surprising that, when we tested English listeners' perception of the Ikalanga /ga^mba-gaba/ stimuli discussed here, their identifications showed some striking similarities to the Ikalanga listeners' performance.

ACKNOWLEDGMENTS

This research was supported in part by NSF Grant BCS-0118684. We are grateful to Rose Letsholo for her collaboration throughout the project, to the University of Botswana for their generosity in making testing facilities available, and to the Ikalangan participants who made this study possible. We also thank members of the phonetics-phonology group at the University of Michigan for helpful comments on earlier versions of this work. Portions of the acoustics findings were presented at the 143rd Meeting of the Acoustical Society of America, June 2002.

REFERENCES

- [1] I. Maddieson, "Prenasalized stops and speech timing," *Journal of the International Phonetic Association*, vol. 19, pp. 57-66, 1989.
- [2] K. Hubbard, "'Prenasalized consonants' and syllable timing: evidence from Runyambo and Luganda," *Phonology*, vol. 12, pp. 235-256, 1995.
- [3] I. Maddieson and P. Ladefoged, "Phonetics of partially nasalized consonants," in *Nasals, Nasalization, and the Velum*, M. Huffman and R. Krakow, Eds., pp. 251-301. New York NY: Academic Press, 1993.
- [4] B. Connell, Personal communication, 2003.
- [5] I. Maddieson, Personal communication, 2003.
- [6] J. Mathangwane, *Ikalanga Phonetics and Phonology: a Synchronic and Diachronic Study*, Stanford CA: CSLI Publications, 1999.
- [7] M. Chen, "Acoustic correlates of English and French nasalized vowels," *Journal of the Acoustical Society of America*, vol. 102, pp. 2360-2370, 1997.
- [8] L. Hyman, "The limits of phonetic determinism in phonology: *NC revisited," in E. Hume and K. Johnson, Eds., *The Role of Speech Perception in Phonology*, New York NY: Academic Press, 2001.
- [9] C. Browman and L. Goldstein, "Towards an articulatory phonology," *Phonology Yearbook*, vol. 3, pp. 219-252, 1986.