CUES TO PANÃRA NASAL-ORAL STOP SEQUENCE PERCEPTION

Myriam Lapierre and Susan Lin

University of California, Berkeley
{myriam.lapierre, susanlin}@berkeley.edu

ABSTRACT

In the Jê language Panãra, speakers produce two auditorily similar consonant sequences, both of which may be transcribed segmentally as a nasal-oral stop sequence [NT], arising from two distinct processes: prenasalization of oral stops /ṼT/ → [ṼNT], and postoralization of nasal stops /NV/ → [NTV]. Previous research on these sequences showed that, despite their segmental similarity, native speakers of Panãra produce these sequences with systematic acoustic differences, most prominently the relative durations of the nasal murmur to the stop closure, but potentially also presence and amplitude of a stop burst and the quality of the nasal murmur. This study demonstrates that the contrast between the two types of NT sequences is perceptually robust among native Panãra speakers, and that all three acoustic cues are utilized by listeners in disambiguating the underlying representation of these two [NT] sequences.

Keywords: speech perception, Panãra, Jê, prenasalized stop, postoralized nasal

1. INTRODUCTION

In Panãra (ISO code: kre), a Jê language spoken by approximately 630 people in the Brazilian Amazon, nasal-oral stop sequences [NT] arise through two distinct processes: prenasalization of underlying oral stops preceded by a nasal vowel (1), and postoralization of underlying nasal stops followed by an oral vowel (2).

(1) /p, t, s, k/ → [mp, nt, ns, ŋk] / Ṽ__
e.g. /põ**p**õ/ → [põ**mp**õ]

(2) /m, n, ɲ, ŋ/ → [mp, nt, ns, ŋk] / __V
e.g. /mĩ**n**ɔ/ → [mĩ**nt**ɔ]

The processes in (1-2) both arise from spreading of the orality or nasality of a vowel onto an adjacent consonant, rather than from a consonant onto an adjacent vowel, as is common in languages spoken in other parts of the world [2, 4, 5, 6, 11]. The postoralization pattern in (2), common in South American languages, can be understood as a form of *shielding*, a phenomenon that renders the contrasts between phonemically oral and nasal vowels maximally distinct [7, 12, 13, 14].

Although the processes in (1-2) ostensibly result in the same segmental surface form [NT], previous research has highlighted significant differences between the two types of [NT]s, in both their articulation and the resulting acoustics [9]. In particular, postoralization (2) appears to be a robust, categorical phonological process, where the nasal murmur and oral stop closure are comparable in duration. In contrast, prenasalization (1) is highly variable between speakers, both in terms of whether it occurs at all and the relative durations of the nasal and oral components. Furthermore, the duration of the nasal component is significantly shorter in prenasalized oral stops than in postoralized nasal stops.

This study is the first to test the perception of allophonic [NT]s, such as those arising from shielding. Our goals are to determine (i) whether native speakers of Panãra can systematically differentiate between postoralized nasal stops and prenasalized oral stops when listening to a given [NT] sequence, and (ii) which acoustic cues listeners rely on in identifying a given [NT] sequence as arising from an underlying /N/ or /T/, and the relative importance of these cues. While previous work has demonstrated differences in the relative durations of nasal murmur and oral stop closure between the two types of [NT] sequences [9], these differences are highly variable between speakers. We thus expect that there may well be other cues that listeners utilize in discriminating these sequences. We hypothesize that these cues may be found in the acoustic qualities of the nasal murmur, or the strength of an oral stop burst. To address our hypotheses, we manipulated speech from one native Panãra speaker (male, 24 years old at time of recording) to have varying acoustic qualities, then presented the audio to other native speakers of Panãra and recorded their judgments regarding the lexical item that best corresponded to each stimulus.

In a similar experiment, Beddor & Onsuwan [3] found that /N-ND/ and /D-ND/ contrasts in Ikalanga (Bantu) are dependent on both *intrinsic* and *extrinsic* cues. Examples of intrinsic cues the authors expected listeners to attend to included the presence of an oral release burst, signaling the presence of /D/, and the presence of nasal murmur, signaling the presence of /N/. Coarticulatory nasalization of vowels surrounding the consonants, on the other hand, is considered an extrinsic cue. For the Ikalangan /N-ND/ contrast, they found that the extrinsic cue was both necessary and sufficient for native listeners to distinguish between the consonants, but the intrinsic cues alone were insufficient [3]. In contrast, for the /D-ND/ contrast, exactly the opposite was true – intrinsic cues alone were sufficient to cue the /D-ND/ contrast, while extrinsic cues were unnecessary.

We predict that Panãra listeners will be able to reconstruct an underlying /N/ or /T/ from a surface [NT] sequence, i.e. that the contrast between prenasalized oral stops and postoralized nasal stops is not (yet) perceptually neutralized. Specifically, we predict that the following acoustic cues will affect listeners' perceptions of [NT] as either /N/ or /T/:

1. The relative duration of the nasal and oral components, where greater ratio of nasal murmur increases /N/ responses;
2. the source of the nasal murmur, where nasal murmur extracted from a plain nasal [N] increases /N/ responses; and
3. the presence or absence of a release burst, where absence of burst increases /N/ responses.

Panãra differs critically from Ikalanga in that oral vs. nasal vowels in Ikalanga are not contrastive [10]. We therefore further predict that our results will differ from Beddor & Onsuwan's [3] in that Panãra listeners will rely more heavily on the intrinsic cues, such as the presence of an oral release burst, over extrinsic cues, such as the quality of the following vowel. To test this, we manipulated the oral vs. nasal quality of the vowel following the [NT] sequence, and predict that listeners will be more likely to perceive [NTṼ] as being underlyingly /TṼ/.

2. METHODS

To test our hypotheses, we prepared eight series of stimuli representing words of the shape /TV/, /TṼ/, /NV/, and /NṼ/, presented in the carrier phrase [kjẽhẽ kasũ X] (*I say the word X*), where X is the target word. This carrier phrase crucially places the target /N/ and /T/s in the context of a preceding nasal vowel, creating the environment required for prenasalization (1). Stimuli were randomized and played for native Panãra speakers in a 4-option forced choice perception task.

2.1. Stimuli

Due to phonotactic and lexical constraints in Panãra, and the desirability of using real Panãra words, we focused on a single minimal-quadruple, presented in (3). The original acoustic items were recorded by one male speaker of Panãra in the Panãra village of Nãnsêpôtiti. He produced five repetitions of each target word embedded in the target carrier phrase.

(3) a. /pa/ → [pa ∼ mpa] (*arm*)
b. /pã/ → [pã ∼ mpã] (*owl*)
c. /ma/ → [mpa] (*liver*)
d. /mã/ → [mã] (*emu*)

The eight series constructed from this speech are summarized in Table 1. We selected one repetition of /pa/ [pa] and one repetition of /mã/ [mã] to serve as the base for series 1 & 3 and 2 & 4, respectively. These acoustic tokens were selected on the basis of maximizing auditory naturalness and consistency in the prosody of the target word and preceding verb. We selected *two* base tokens rather than one in order to test our hypothesis that the quality of the nasal murmur would affect listeners' perceptions. In Series 1 & 3, we spliced in nasal murmur cut from one designated /ma/ [mpa] production, and in Series 3 & 4, the nasal murmur comes from the /mã/ [mã] already present in the base sentence.

To test our hypothesis that listeners would differentiate between underlying /N/ and /T/ based on differences in duration of nasal vs. oral components, we varied the relative duration of the nasal stop murmur and oral stop closure. First, we normalized the target consonants [p] and [m] in these two selected base recordings to approximately 160ms, by cutting or repeating sections of the waveform. For Series 1 & 3, we replaced 40ms of silence at the beginning of the oral stop duration with 40ms of nasal murmur per step. For Series 2 & 4, we replaced 30ms of nasal murmur at the end of the nasal stop duration with 30ms of silence per step. This procedure resulted in uneven numbers of steps – Series 1 & 3 have 5 steps while Series 2 & 4 have 7 – but the end points are consistent: 0ms nasal murmur and 160ms stop closure vs. 160ms nasal murmur and 0ms stop closure.

To test our hypothesis that nasality or orality of the vowel following the target [NT] would affect listeners' perceptions, we included both oral and nasal

Series	Vowel	Murmur	Closure	Burst
1a	[a]	[mpa]	[pa]	no
1b				yes
2a		[mã]	silence	no
2b				yes
3a	[ã]	[mpa]	[pa]	no
3b				yes
4a		[mã]	silence	no
4b				yes

Table 1: Acoustic parameters of the 8 series, including quality of following vowel, source of nasal murmur, stop closure, and presence of burst.

vowels in the manipulation. As outlined in Table 1, Series 1 and 2 have the oral vowel [a] from [pa], and Series 3 and 4 have the nasal vowel [ã] from [mã].

Finally, to test the effect of oral stop bursts in underlying /N/ or /T/ perception, from each of the four Series described above, we created two series, such that every stimulus item in (a) Series contains no oral stop burst between the end of the consonant and the onset of the following vowel, and every stimulus item in (b) Series contains an audible stop burst.

2.2. Participants and procedures

Perception data was collected in the village of Nãnsêpôtiti from 36 native listeners of Panãra (17 male) between the ages of 16 and 40 (mean age=25). The intent to conduct an experiment was presented to community leaders who spread the word. Individuals who were interested presented themselves at the researcher's house to participate.

During the task, participants were presented with laminated physical printouts of the four photos shown in Figure 1, each representing one of the target words in (3). Prior to commencing the task, participants were asked to name each picture. The task did not begin until they had pronounced each one of the target words to confirm their knowledge of the lexical items and intended phonological form.

Figure 1: Visual stimuli.

Participants listened to stimuli through a pair of semi-open circumaural headphones (AKG K240 studio) one at a time, and pointed to the image representing the word they heard. They were free to ask the researcher to repeat the recording as many times as they wished before providing an answer. Each stimulus item was repeated 5 times, in semi-randomized blocks. Participants were compensated at the conclusion of the experiment[1].

Some participants' responses appeared to be random, and not conditioned by the acoustic manipulations present in the stimuli. In order to determine which participants fell into this category, for each stimulus item, we defined a set of 'error' responses to be any response that made up less than 10% of the total responses for that stimulus. Then, for each participant, we calculated his or her proportion of 'errors', and the full set of data from any participant with greater than 10% 'error' rate was excluded. This procedure resulted in the removal of six female participants, whose 'error' rates ranged between 11%-38%. The final pool of participants whose data were analyzed was 30 participants, 13 female and 17 male.

3. RESULTS & DISCUSSION

3.1. Relative duration & source of nasal murmur

As predicted, listeners' perception of the stimuli varied by the relative duration of the nasal murmur compared to the stop closure. As shown in Figure 2, whether the following vowel was oral (triangles) or nasal (circles), and whether the nasal murmur was derived from a plain nasal [m] (dashed lines) or postoralized [mp] (solid lines), as the relative duration of the nasal murmur in the [NT] sequence increased, the proportion of /m/ responses rose.

In Series 1 & 2, the preceding vowel was nasal and the following vowel was oral (/Ṽ_V/), resulting in a licit phonotactic environment for *both* prenasalization of /pa/ and /pã/, and postoralization of /ma/. It is thus unsurprising that responses vary by nasal murmur duration to the extent that they do, as that duration appears to be a consistent distinction made by native Panãra speakers during the production of

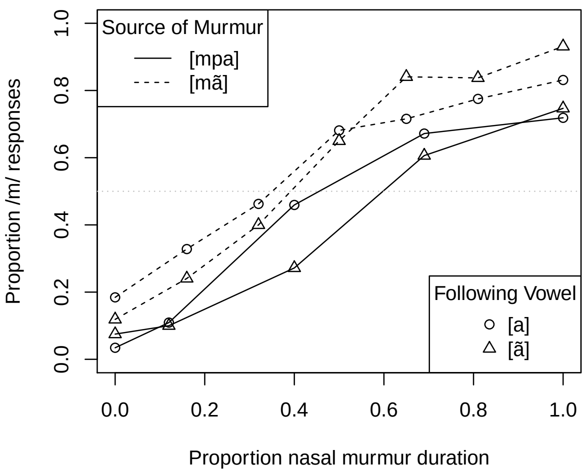

Figure 2: Proportion of /m/ responses, by source of nasal murmur and following vowel.

the /m/ vs. /p/ contrast in this environment [9].

On the other hand, in Series 3 & 4, both preceding and following vowels were nasal (/Ṽ_Ṽ/), resulting in a phonotactically *illicit* environment for postoralization. Thus when the duration of nasal murmur is long and the source of murmur is plain [m], the proportion of /mã/ responses is very high and the proportion of /pã/ responses correspondingly low. It is interesting to note that when the source of nasal murmur is a postoralized /ma/ [mpa], listeners cross over later in the series than when the source is /mã/ [mã], and listeners do not perceive the 100% nasal murmur endpoint of Series 3 as /m/ 100% of the time. This modulating effect of the source of the nasal murmur is significant regardless of whether the following vowel is nasal ($p = 0.0209$) or oral ($p = 0.0491$). Here and elsewhere, comparisons were made using a mixed effects logistic regression model, with speaker as a random factor [1].

3.2. Stop burst & following vowel nasality

It is also not very surprising that the presence of a burst resulted in a greater proportion of /p/ responses in all four series. As shown in Figure 3, whenever a burst was present (black lines), listeners required a greater proportion of nasal murmur to perceive /m/ compared to when the burst was absent (gray lines). When the origin of the nasal murmur was from postoralized /ma/ [mpa], the effect of the burst was significant but moderate (Series 1: $p < 0.0001$; Series 3: $p < 0.0001$), and roughly equivalent regardless of the nasality on the following vowel. In contrast, when the origin of the nasal murmur was /mã/ [mã] (Series 2 and 4), the effect of the burst interacted with quality of the following vowel. In particular, this effect is not significant in Series 2 ($p = 0.0930$) where the following vowel is oral, and is pronounced in Series 4 ($p = 0.0006$), where it is nasal.

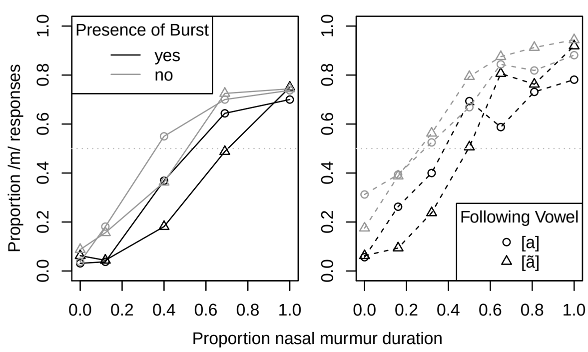

Figure 3: Proportion of /m/ responses, by following vowel quality and burst. Nasal murmur from postoralized /m/ (left) vs. plain /m/ (right).

One interpretation of this result is that when the murmur is from /mã/ [mã] but is followed by an oral vowel (Series 2), the presence or absence of an audible stop burst has a reduced effect because ,even if the listener perceives the sequence as [mp], listeners are likely to attribute any evidence of [p], including a burst, to postoralization. On the other hand, when it is followed by a nasal vowel (Series 4), the presence or absence of a burst becomes the primary cue for the presence of an oral stop, and if that stop is perceived, the only phonotactically licit choice for listeners is prenasalized /pã/.

4. CONCLUSION

This paper presented the results of a perception experiment testing the cues to which native Panãra listeners attend in identifying a surface [NT] sequence as arising from a prenasalized oral stop /T/ (1) or a postoralized nasal stop /N/ (2). Our results suggest that listeners' responses are dependent on all manipulated acoustic cues. Longer nasal murmurs resulted in greater /N/ responses, as did the lack of an oral release burst and the presence of nasal murmur derived from /mã/ [mã]. The oral vs. nasal quality of the following vowel interacted with these findings: when the target consonant was followed by a nasal vowel, the presence of an audible burst appeared to be a more important cue to identifying the underlying representation than the quality of the nasal murmur.

The results of our study are complementary to those of Beddor & Onsuwan [3]. Given that Panãra vowels contrast nasality while Ikalanga vowels do not, we predicted that vowel quality would *not* serve as an extrinsic cue for Panãra speakers. Furthermore, voiceless stops are associated with a more prominent release burst and a relatively longer oral component than voiced stops [15], and we therefore predicted that the oral release burst would be a more prominent perceptual cue for Panãra speakers than for Ikalanga speakers. These predictions are borne out by our results, which suggest that the intrinsic cues (e.g. duration of nasal murmur or presence of oral release burst) are vital to /N/ vs. /T/ perception, while the extrinsic cue of vowel nasality is of consequence only to the extent that listeners use it to rule out phonotactically illicit sequences. These results are consistent with predictions by Kaplan [8] and Stanton [12].

5. ACKNOWLEDGEMENTS

We are sincerely grateful to the Panãra community for allowing us to study their language, especially to Tepakriti Panãra, whose voice was used in preparing the stimuli. We thank Julia Schweers for support in collecting the data. This research was supported by two Robert Oswalt Endangered Language Grants.

6. REFERENCES

[1] Bates, D., Mächler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.

[2] Beddor, P. S. 2009. A coarticulatory path to sound change. *Language* 84(4), 785–821.

[3] Beddor, P. S., Onsuwan, C. 2003. Perception of prenasalized stops. *Proceedings of the 15th International Congress of Phonetic Sciences* 407–410.

[4] Cohn, A. C. 1990. Phonetic and phonological rules of nasalization. *UCLA Working Papers in Phonetics* 76.

[5] Delvaux, V., Demolin, D., Harmegnies, B., Soquet, A. 2008. The aerodynamics of nasalization in french. *Journal of Phonetics* 36(4), 578–606.

[6] Desmeules-Trudel, F., Brunelle, M. 2018. Phonotactic restrictions condition the realization of vowel nasality and nasal coarticulation: Duration and airflow measurements in Québécois French and Brazilian Portuguese. *Journal of Phonetics* 69, 43–61.

[7] Herbert, R. K. 1986. *Language Universals, Markedness Theory, and Natural Phonetic Processes*. Berlin: Mouton de Gruyter.

[8] Kaplan, A. 2008. Markedness and phonetic grounding in nasal-stop clusters. Qualifying Paper, Department of Linguistics; University of California, Santa Cruz.

[9] Lapierre, M., Lin, S. 2018. Patterns of nasal coarticulation in Panãra. Poster presented at the 16th Conference on Laboratory Phonology.

[10] Mathangwane, J. 1996. *Phonetics and phonology of Ikalanga: a diachronic and synchronic study*. PhD thesis University of California, Berkeley.

[11] Scarborough, R., Zellou, G., Mirzayan, A., Rood, D. S. 2015. Phonetic and phonological patterns of nasality in Lakota vowels. *Journal of the International Phonetic Association* 45(3), 289–309.

[12] Stanton, J. 2017. *Constraints on the distribution of nasal-stop sequences: an argument for contrast*. PhD thesis Massachusetts Institute of Technology.

[13] Storto, L. 1999. *Aspects of a Karitiana grammar*. PhD thesis Massachusetts Institute of Technology.

[14] Wetzels, W. L., Nevins, A. 2018. Prenasalized and postoralized consonants: The diverse functions of enhancement. *Language* 94(4).

[15] Zue, V. W. 1976. *Acoustic characteristics of stop consonants: A controlled study*. PhD thesis Massachusetts Institute of Technology.

[1] Participants received 83 grams of Czech beads as compensation for taking part in the experiment. Note that this type of remuneration was chosen at the request of the community. The economic system in use inside the Panará Indigenous Land is mostly dependent on trade, and monetary transactions are virtually absent inside the villages. Czech beads are among the most valued goods, as they are used to create ornaments worn during village celebrations and festivities.