Phonetics of modal and breathy nasals in Xitsonga

Crous M. Hlungwani[1], Seunghun J. Lee[1,2]
University of Venda[1], International Christian University[2]
Crous.Hlungwane@univen.ac.za[1], seunghun@icu.ac.jp[2]

ABSTRACT

Xitsonga is a southern Bantu language with a phonation contrast in nasals. This paper reports results from acoustic and articulatory data collected in Thohoyandou, Limpopo. This study is the first to examine Xitsonga breathiness with Electro-glottography (EGG). The EGG analyses of 6 Xitsonga speakers show that the contrast varies according to tonal contexts. In the high tone context, modal nasals demonstrated lower open quotient than breathy nasals. In the toneless context, however, the contrast was neutralized. Moreover, female speakers have a breathier voice than male speakers. As in [8], f0 values were consistently lower in breathy tokens.

Keywords: Xitsonga, nasal, phonation, breathy, Electroglottograph (EGG), perception

1. INTRODUCTION

Xitsonga (S53, [4], also known as Shangaan), a southern Bantu language, is one of the eleven official languages spoken in Limpopo, South Africa. Xitsonga has an extensive dictionary [2] and also a grammar book [1]. Xitsonga is a two-tone language (high and toneless) with a contrast between modal and breathy phonation in the nasals. This phonation contrast in Xitsonga has been investigated in [8] where inter- and intra-speaker variation were reported. This paper explores an articulatory study of the nasal phonation type by examining data from electroglottograph (EGG) recordings. Examples of Xitsonga phonation contrast in IPA are shown in (1), where breathy nasals have a diaeresis under the nasal:

(1) Modal vs. breathy nasal

[mafundza] 'respect'	[m̤aŋgu] 'choose'
[nalá] 'enemy'	[n̤amu] 'neck'
[ŋaŋga] 'doctor'	[ŋ̤ʷeti] 'month'

The breathy phonation is interesting phonologically as well, because when coupled with a low tone, the breathy consonants block the spreading of high tone; a general tonal process in Xitsonga. The breathy nasals are tone-lowering and thus they are called depressors. Acoustic studies on depressors are uncommon, with some exceptions ([6] on Tsua) that report item variability, especially when high tone is present. [8] is the only acoustic study of breathy nasals in Xitsonga. They show that breathy nasals have large perturbations in f0, and large speaker variation by gender is found. [8] concludes that spectral tilt is an indicator for distinguishing breathy from modal nasals. Complementing [8], the current study explores how the two phonation types of nasal categories are produced articulatorily and distinguished perceptually.

2. METHOD

The data reported is based on the fieldwork in Thohoyandou, Limpopo, South Africa, which was conducted in November 2017.

2.1. Speakers

Thirteen native speakers of Xitsonga participated in the recording session. They were all university students majoring in the Xitsonga language. All the speakers spoke English in addition to Xitsonga. Some speakers had basic knowledge of Tshivenda. The age ranged from 21 years old to 25 years old. Before the recording session, consent forms and demographic questionnaires were collected from each speaker. Each participant was compensated for their time (100 South African Rand). This paper reports results from six speakers.

2.2. Recording: EGG and Acoustics

Within each recording session, each speaker read target words in a toneless frame sentence *ni tirhisa X kan'we* 'I use X again' and in a high-toned frame sentence *vá tírhísá X kan'we* 'They use X again'. Each of these tokens was recorded five times. Speakers wore a two-headed electrode around their neck and a head-worn microphone (an XLR Shure WH-30) that were connected to the Electroglottograph (EG2-PCX2, Glottal Enterprises) device. The output analog signals from the microphone and the electrodes were captured by an external sound card that was directly connected to a Macintosh computer via a Roland USB Audio Interface Rubix 24. Stereo recordings of these signals were made using Praat. After the recording session, participants were asked to validate a list of stimuli and check whether they use a particular word with other Xitsonga speakers. The order of the target words was randomized, and the speakers repeated the list three times. The stimuli were presented in Latin script using PowerPoint on a second Macintosh computer. The target of the recording session included 3 breathy nasals and 3 modal nasals, although the current analysis pools data from contexts across different places of articulation.

2.3. Perception test

Six listeners of Xitsonga participated in the perception test in November 2018. Participants were asked to judge whether a word in a frame sentence begins with a modal nasal or a breathy nasal (spelled with 'h' after a nasal. Each participant responded to 200 trials. The perception test was run using Superlab 5 on a Macintosh Computer with two keys on the keyboard as input method. Participants wore Sennheiser noise reducing headphones.

2.4. EGG analysis

The EGG data set consists of 180 tokens (6 speakers * 6 targets * 5 repetitions), which were analyzed using Eggnog v.0.3 [3]. This function reads EGG recordings from a directory and returns a file that can be imported in R for further analysis. Figure 1 provides an example of an Eggnog visualization where an EGG signal is shown along with F0 and open quotient (Oq) calculated from the EGG signal. Eggnog is freely available from [3], and can be used in conjunction with a Matlab license and its signal processing toolbox.

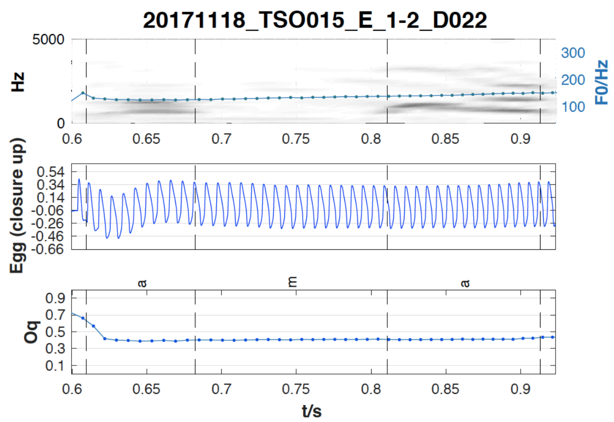

Figure 1: Results from Eggnog of labial nasal with modal phonation in a toneless context (male)

2.5. Acoustic analysis

Data from six Xitsonga speakers also collected in November 2017 were processed for further acoustic analysis using VoiceSauce [7]. F0 was extracted using the Straight algorithm at a 10 ms interval. One spectral tilt measure (H1-H2) as well as Harmonic-to-Noise ratio at various frequencies were also obtained using VoiceSauce.

3. RESULTS

Results in this section focus on the production of one male speaker and one female speaker.

3.1. EGG results

Results from EGG data show that the phonation distinction is not as prominent as expected in the production of Xitsonga speakers. In the toneless context, the production of breathy items (Figure 2) was similar to the modal phonation (compare with Figure 1). The breathy nasal has a slight increase of Oq near the midpoint, but it was not significant. Both male and female speakers show the same kind of pattern.

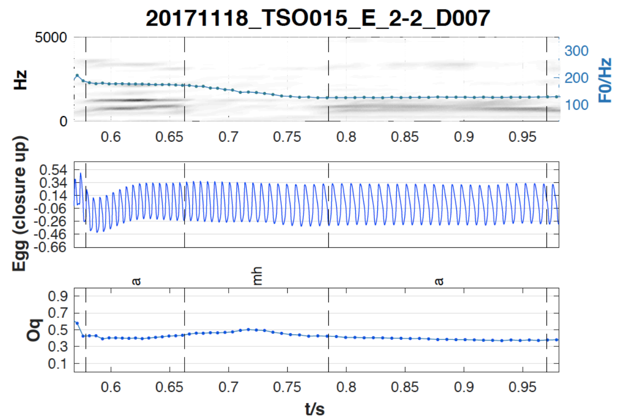

Figure 2: Eggnog results of labial nasal with breathy phonation in a toneless context (male)

In the high tone context, there is a gender difference. The male speaker demonstrates that the breathy phonation is not observed as the Oq is nearly 50%, suggesting modal phonation (Figures 3 and 4). In breathy phonation, the male speaker shows a slight increase in Oq in the middle of the nasal.

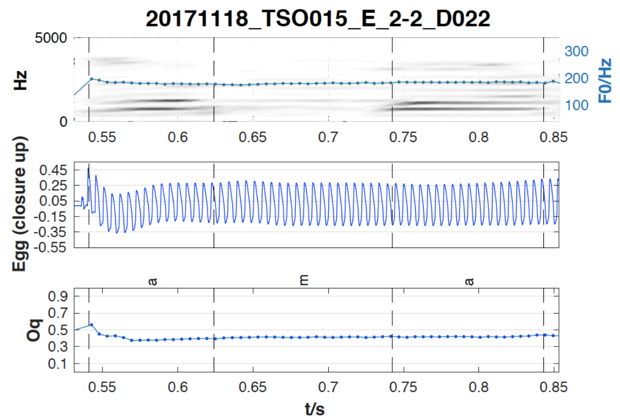

Figure 3: Eggnog results of labial nasal with modal phonation in high tone context (male)

The female speaker in the high tone context has overall higher Oq in both modal and breathy phonation, suggesting that all tokens were produced with breathier phonation (Figures 5 and 6).

In sum, the articulation results agree with [8] in that there is a gender difference found in the production of breathy nasals. The absence of contrast in the articulation of male speakers begs the question of how Xitsonga speakers perceive this contrast.

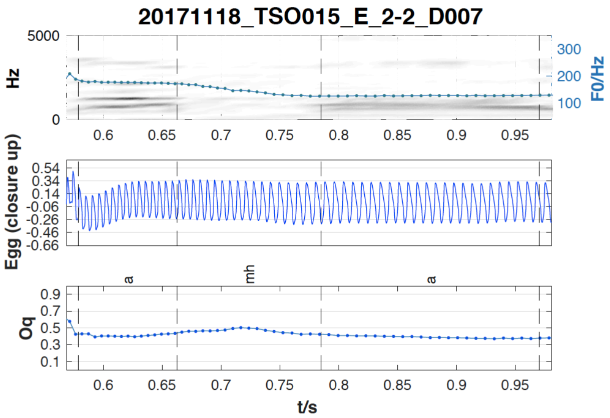

Figure 4: Eggnog results of labial nasal with breathy phonation in high tone context (male)

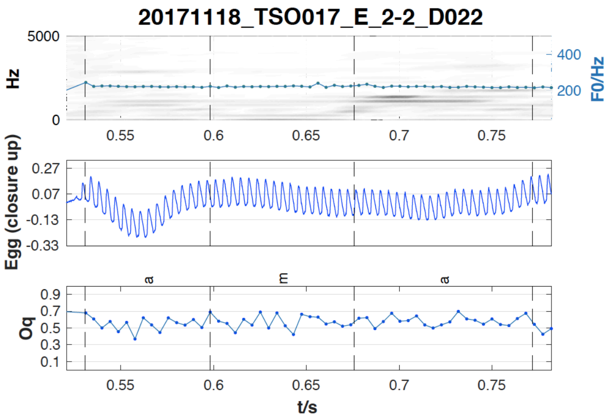

Figure 5: Eggnog results of labial nasal with modal phonation in high tone context (female)

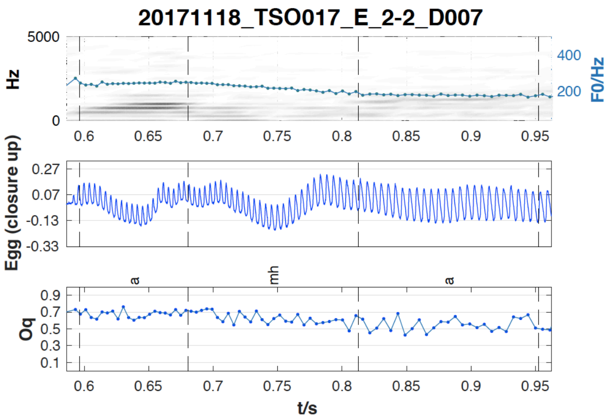

Figure 6: Eggnog results of labial nasal with breathy phonation in high tone context (female)

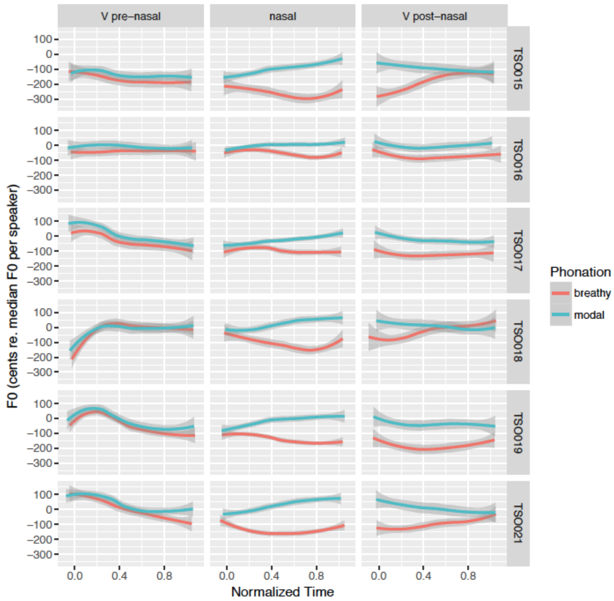

Figure 7: F0 plots in the toneless context. From left to right, each panel represents a preceding vowel, target nasal and a following vowel.

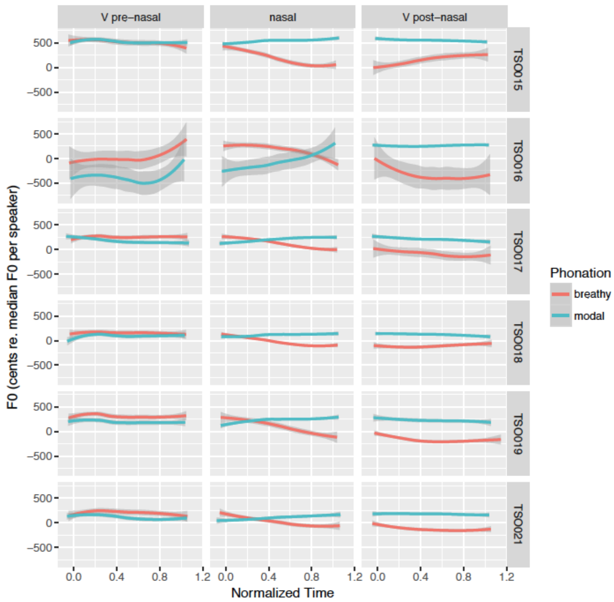

Figure 8: F0 plots in the high tone context. From left to right, each panel represents a preceding vowel, target nasal and a following vowel.

3.2. Acoustic results

Acoustic analysis for the nasal phonation was conducted with VoiceSauce [7]. In the toneless context, f0 of breathy nasals becomes lower throughout the nasal compared to that of modal nasals as in Figure 7. This F0 difference continues throughout the following vowel. In the high tone context in Figure 8, the f0 plots show individual variation. For most speakers, the lowering of f0 induced in the breathy nasal only begins after the production of the nasal. The f0 lowering in the following vowel is more visible than in the nasals themselves, as coarticulation of H tone of a preceding vowel affects f0.

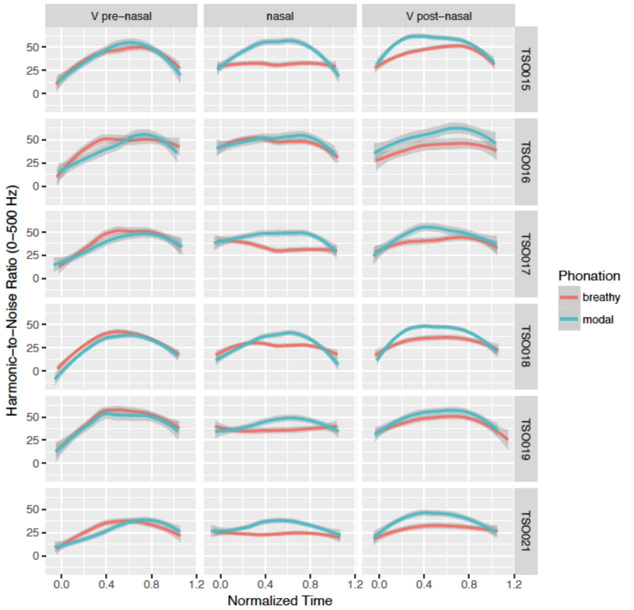

Figure 9: Harmonic-to-Noise ratio at 0 ~ 500 Hz in the H tone context.

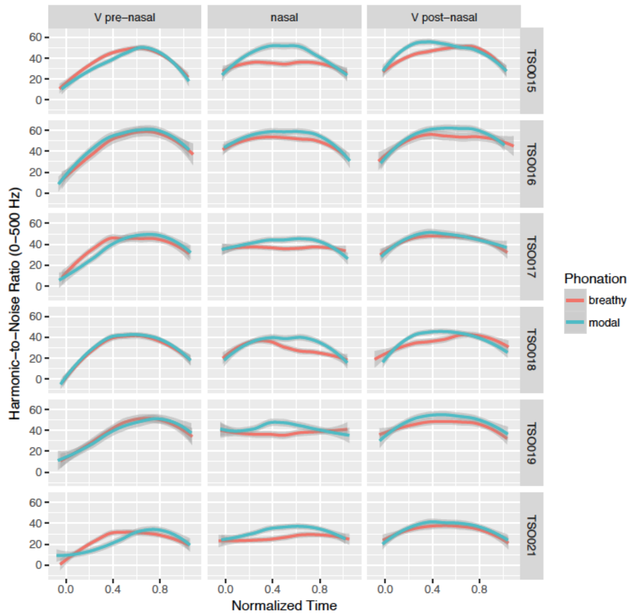

Figure 10: Harmonic-to-Noise ratio at 0 ~ 500 Hz in the toneless context.

Spectral tilt measurements did not produce a great difference between modal and breathy nasals. Harmonic-to-noise ratio obtained at various frequency levels show that speakers have a higher level near the midpoint of a modal nasal, both in the toneless and the high tone context. Figures 9 and 10 show the Harmonic-to-Noise ratio at 0 ~ 500 Hz in both contexts.

3.3. Perception results

Results of the perception experiment in Figure 11 show that Xitsonga speakers can identify the breathy nasal from the modal nasal with high accuracy, confirming the presence of the phonation contrast.

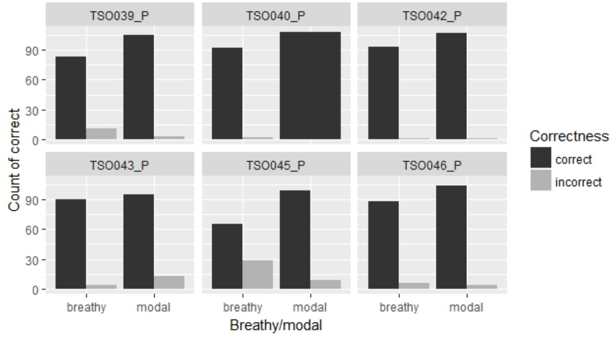

Figure 11: Perception results for Xitsonga speakers. Black bars represent correct responses for each category[1].

4. CONCLUSION

The current paper set out to explore how Xitsonga speakers realize a contrast between breathy nasals and modal nasals, in particular because this phonation difference has phonological consequences [1, 5]; breathy nasals block high tone spreading but modal nasals do not block high tone spreading. Results from the articulatory study show that all nasals by the male speaker were produced as modal-like. Breathy nasals produced by the female speaker were breathier, but so were modals, indicating a general tendency for breathiness, rather than any phonation contrast, in the female speaker.

Articulatory results show unexpected results with modal-like quality (low Oq), but acoustic analysis show that pitch lowering (i.e. slower vibration) is a main factor for distinguishing modal from breathy nasals. In addition, the frame sentence context could have affected the production of breathy nasals because a speaker would have had to change the patterns of vocal fold vibration on the fly. Data in isolation could reveal different results, which we will leave for future research.

5. ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Number 16K02641 to the second author. An earlier version of this study appears in the Proceedings of the 3rd Asian Junior Linguists Conference. We thank Julian Villegas, Jeremy Perkins, William Bennett, Aaron Braver, and the audience from CIL and AJL3, and three anonymous reviewers for their comments.

[1] TSO040 provided correct responses for all the modal nasals correct. Thus, there is no gray bar in the panel.

6. REFERENCES

[1] Baumbach, E. 1987. *Analytical Tsonga*, UNISA.
[2] Cuenod, R. 1997. *Xitsonga dictionary*. Sasavona Books.
[3] J. Villegas, "Eggnog: ElectroGlottoGraphy Not Obligatory Guidance," 2018. Version 0.3. Available from http://onkyo.u-aizu.ac.jp/software/eggnog/
[4] Guthrie, M. 1967–1971. *Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages*. Farnborough: Gregg Press.
[5] Lee, S. J. (2008) Consonant-tone interaction in Optimality Theory, Ph.D. Dissertation, Rutgers University.
[6] Mathes, T. 2014. Extreme Tonal Depressor Effects in Khoisan: Evidence from Tsua. *Proc. 4th TAL* Nijmegen.
[7] Shue, Y.-L., Keating, P., Vicenik C., Yu K. 2011. VoiceSauce: A program for voice analysis, *Proc. 17th ICPhS* Hong Kong, 1846-1849.
[8] Traill, A., Jackson M. 1988 Speaker variation and phonation type in Tsonga nasals. *Journal of Phonetics* 16: 385-400.