An Electroglottograph Study of Xitsonga nasals: an exploratory study

Seunghun J. Lee
International Christian University & University of Venda

1 Introduction

Field research on the sound systems of a language has traditionally focused on describing sounds that are contrastive. Reflecting this trend, grammars of understudied languages have a phonology section with the necessary information to understand the basics of the language, but often lacks detailed phonetic descriptions of the sounds. It is not uncommon to find grammars that discontinue marking tone in a tonal language when reporting on other areas of grammar. While documenting words and sentences is important, phonetic documentation of plosives, fricatives, as well as other suprasegmentals, is equally important, especially when the language is understudied or endangered. Phonetic documentation not only describes the present status of sounds, but it also provides insight into the role of dialect, interspeaker, or intraspeaker variations. Indirectly, phonetic studies also offer insight into how orthography influences pronunciation patterns.

Phonetic research in language documentation often has been conducted at laboratories away from the natural environment where the language is spoken. Notably, work by Ladefoged and Maddieson was an exception. Currently, however, more phoneticians are collecting data in the field, owing to increased portability of equipment that can be used for doing phonetic investigation. Also, free software such as Praat and R, as well as portable articulatory equipment such as electroglottograph (EGG) and ultrasound, make it more feasible for phoneticians to conduct research in the field. Phonetic findings of field data now open ways to confirm sounds that are previously collected using impressionistic methods.Thus, phonetics and field linguistics need closer working relationships in order to offer a full description of a sound system of an understudied or endangered language.

This study on Xitsonga nasals focuses on using EGG, a method which allows indirect access to the movement of the vocal folds, which are not otherwise readibly visible. EGG measures patterns of the opening and closing of the vocal folds by placing two electrodes around the neck of a participant. In the current setup, EGG data is recorded using Praat via a Roland USB Audio Interface Rubix 24 that takes analog audio and EGG signals directly from the EGG machine (EG2-PCX2 from Glottal Enterprises). Audio signals are recorded using a XLR Shure WH-30 head-word microphone. EGG recordings are then processed using Eggnog, a matlab-based visulation tool developed by Dr. Julián Villegas as shown in figure 2.

Figure 1: From left to right. Two electrodes, EGG machine and Roland Audio Interface

Eggnog represents EGG data with three vertical panels. The mid panel is the EGG signal indicating vocal fold closure in the upward direction. Even if the recordings are closure-down, Eggnog will automatically transform the signals to closure-up. The top panel and the bottom panel represent pitch (in f0) and Oq (open quotient), respectively. These two values are directly calculated from the recorded EGG signals. Open quotient measures the percentage of opening of the glottis within a glottal cycle. The larger the open quotient, the breathier the sound. Sounds with breathy phonation have open quotients of about 70%, while sounds with modal phonation have open quotients of about 50%.

Proceedings of *AJL 3*

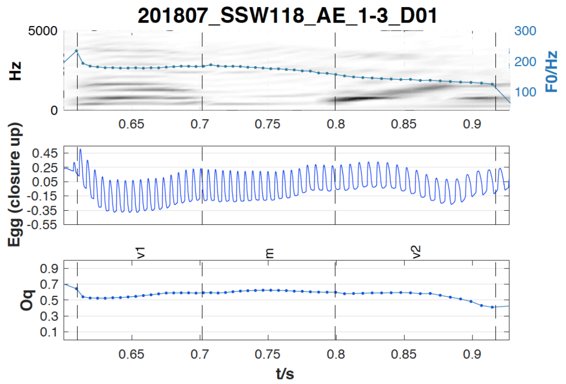

Figure 2: Results of Eggnog (http://onkyo.u-aizu.ac.jp/index.php/software/eggnog/)

2 Xitsonga nasals

Xitsonga is a southern Bantu language (S53, Guthrie 1957-1961) spoken by about 2 million people in South Africa, Mozambique and Zimbabwe. This section is based on the South African variety of Xitsonga spoken in the northeastern part of Limpopo. As one of the eleven official languages, Xitsonga speakers have access to education in their own mother tongue. At the University of Venda, Xitsonga students can major in Xitsonga. The language can also be used when writing graduate level theses, Master's and PhD's. Even so, not much phonetic research is available on Xitsonga.

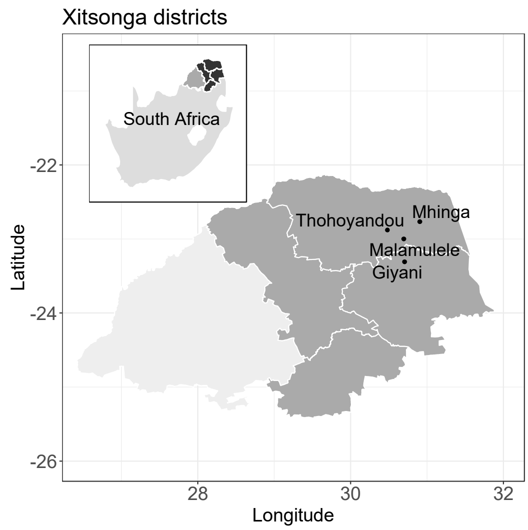

Figure 3: Xitsonga speaking areas

2.1 *Traill and Jackson (1998)*

Traill and Jackson (1988) report that breathy nasals in Xitsonga have larger perturbations and more speaker variation than the modal counterparts. They constructed stimuli based on three places of articulation ([m, n, ŋ]), as well as tonal contexts (high versus low). Stimuli were validated with each speaker ensuring that they produced a phonation contrast between breathy nasals and modal nasals. Using the frame sentence *ndziri X* "I say X", 15 participants recorded the stimuli list two times. Due to the low frequency of lexical items with breathy nasals, minimal pairs or near-minimal pairs could not be included. Their findings show that Xitsonga speakers produce breathy nasals in terms of where the nasality start. The nasal airflow difference for breathy nasals vs. modal nasals is (i) from the beginning of the nasal until the beginning of the following vowel, or (ii) from mid point of a nasal to the midpoint of the following vowel. In addition, Traill and Jackson report that considerable speaker variation based on gender is found; also, spectral tilt is a reliable measure for understanding breathy nasals in Xitonsga.

Aerodynamic measurements and acoustic analyses in Traill and Jackson (1998) were in line with the descriptions of Xitsonga by Louw (1968) and Baumbach (1987), which reported that breathy nasals have an effect of lowering tone, which persists throughout the syllable. Aerodynamic measurements suggest how the status of glottis is producing breathy nasals, but it doesn't tell how vocal cords vibrate during the production of breathy nasals.

An EGG study can fill this gap because compared to modal nasals, breathy nasals have flapping movements of vocal cords that result in an opening of the vocal folds for a longer period of time. The following section reports on investigating nasals in Xitsonga using EGG. Using EGG to investigate breathy nasals is especially useful, since EGG provides data to distinguish phonation types.

3 Current study

3.1 *Data collection and processing* Xitsonga data was collected in November 2017 at the University of Venda located in Thohoyandou, Limpopo. None of the six speakers lived outside of the Limpopo region for any extended period. This section is based on simultaneous acoustic and EGG recordings from one male and one female speaker. All tokens were produced in a toneless and a H tone context, where speakers read target words in a frame sentence from powerpoint slides. The toneless context was *ni tirhisa X kan'we* "I use X again", and the H tone context was *<u>vá</u> tírhísá X kan'we* "They use X again", in which the subject prefix *vá-* introduces a H tone. The list of stimuli is as follows:

	Xitsonga	IPA	Xitsonga	IPA	Xitsonga	IPA
breathy	mhala	[m̤ala]	mhangu	[m̤aŋga]	mhawurí	[m̤awurí]
	nhanga	[n̤aŋa]	nhamu	[n̤amu]	nhómpfú	[n̤ómpfú]
	n'hwari	[ŋ̈wari]	nun'hwela	[nuŋ̈ʷela]	nhweti	[ŋ̈ʷeti]
modal	manana	[manana]	mafundza	[mafundza]	mahlanta	[małanta]
	nala	[nala]	nandza	[nandza]	námbu	[námbu]
	n'an'ana	[ŋaŋana]	n'anga	[ŋanga]	n'wala	[ŋʷala]

Table 1: Xitsonga stimuli list used in this study.

3.2 *Results: Toneless context* Based on the EGG data, open quotient (Oq) is calculated in order to understand the laryngeal setting, i.e., normal, breathy, or creaky. It is calculated from the ratio of the opening phase and the closing phase of a laryngeal cycle. When duration of the opening phase and the closing phase is similar (Oq = 0.5), the resulting sound is said to be a modal voice; when the opening is longer (Oq > 0.7), it is said to be a breathy voice, (whereas creaky voice has a shorter opening phase (Oq < 0.3)).

The results of the Xitsonga EGG data show patterns that were unexpected, and different from what was described in Traill and Jackson (1988). The EGG data we report here that show no differences in Oq for the modal and breathy phonations. The results of breathy and modal nasals processed with Eggnog are shown in figures 4 to 6 (for each place of articulation, figure 4 for labial, figure 5 for coronal, and figure 6 for dorsal). In all contexts produced by the male speaker, the open quotient (Oq) value is around 0.6, indicating that the breathy tokens and modal tokens are not distinguished by Oq.

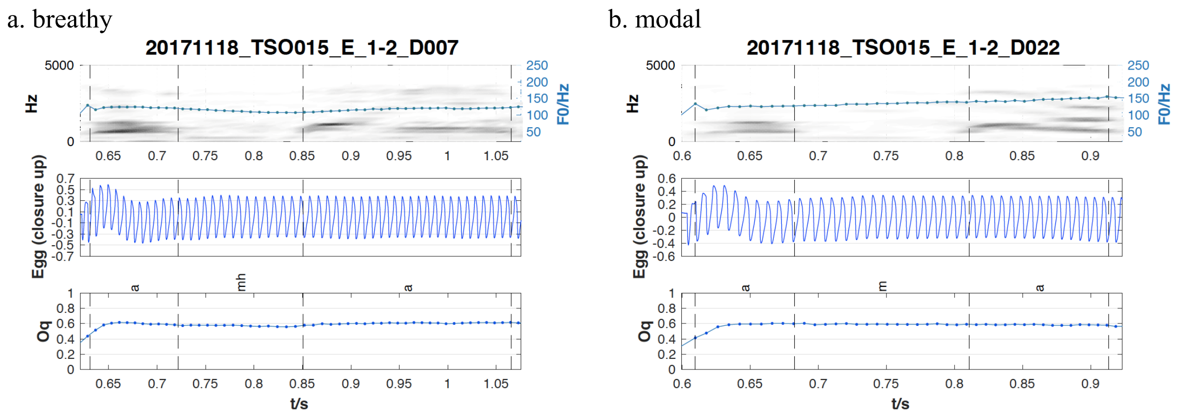

Figure 4: Breathy and modal labial nasals in a toneless context by a male speaker

a. breathy b. modal

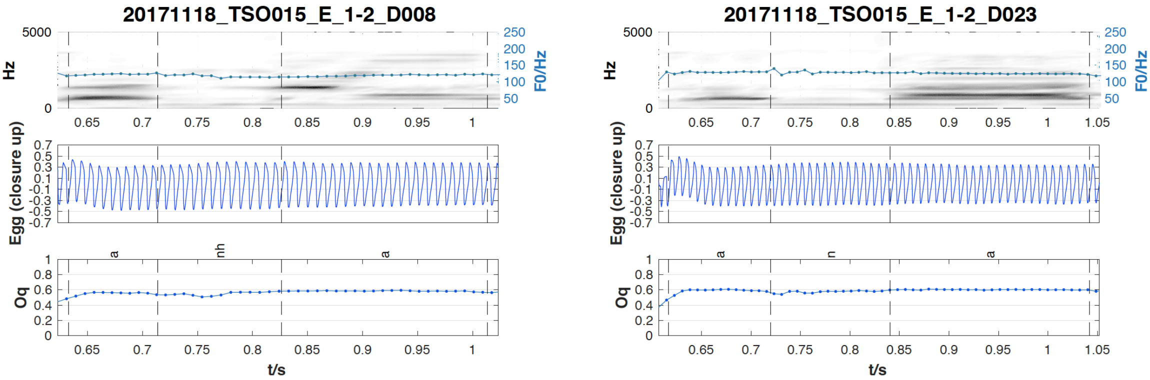

Figure 5: Breathy and modal coronal nasals in a toneless context by a male speaker

a. breathy b. modal

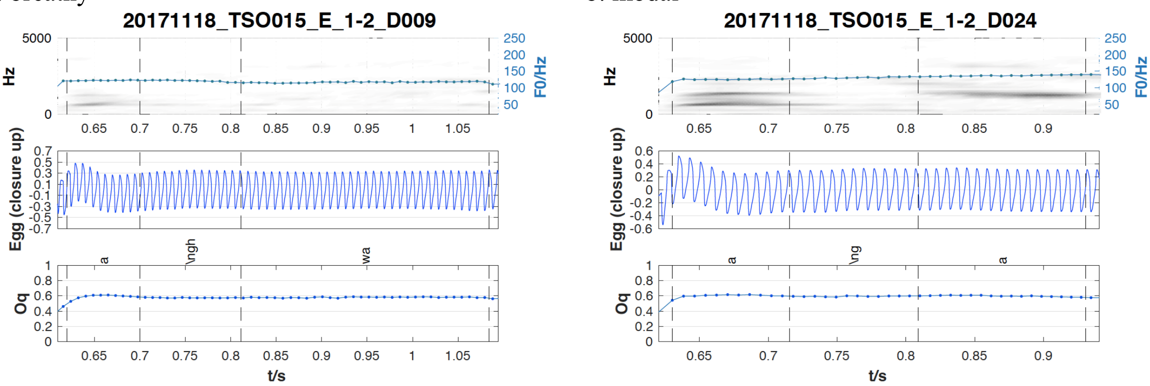

Figure 6: Breathy and modal dorsal nasals in a toneless context by a male speaker

3.3 *Results: H tone context* In the high tone context, the EGG analysis shows that the male speaker has an overall breathy-like voice with the Oq value around 0.6. However, the first half of breathy targets shows a dip around 0.7 s in the Oq value, suggesting that the nasal sound is produced modal or less breathy. This pattern in the breathy nasals is observed across all places of articulation (figure 7 for labial, figure 8 for coronal and figure 9 for dorsal). Modal voice, on the other hand, has consistent Oq values throughout the nasal consonant. At this point, it is not possible to conclude whether the dip for the breathy phonation in Oq is the cue used by native speakers of Xitsonga to distinguish breathy from modal phonation.

a. breathy b. modal

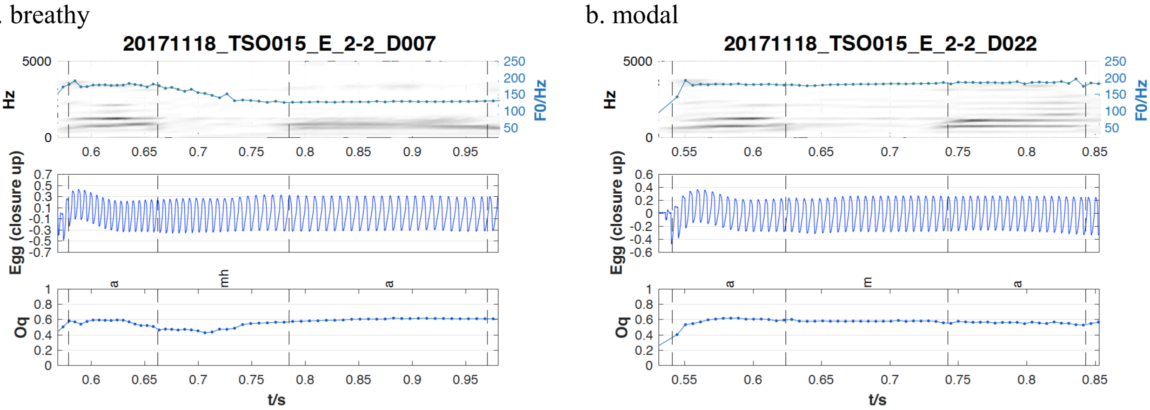

Figure 7: Breathy and modal labial nasals in a H tone context by a male speaker

a. breathy b. modal

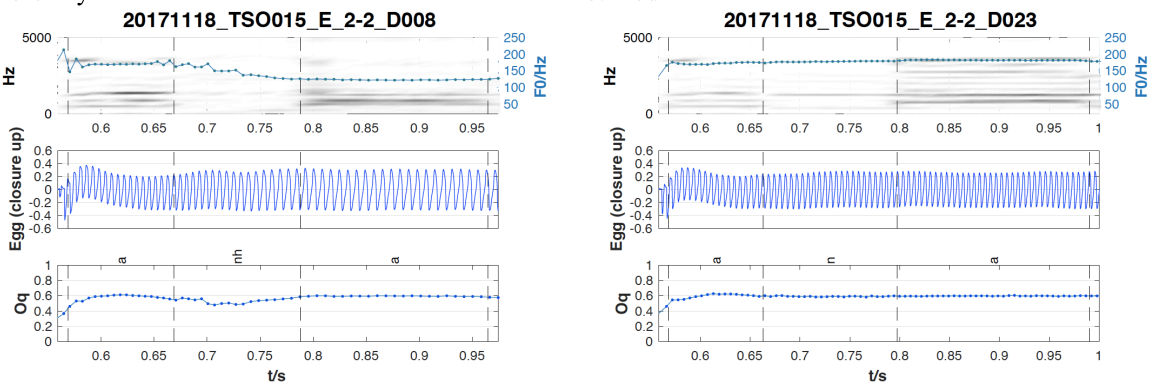

Figure 8: Breathy and modal coronal nasals in a H tone context by a male speaker

a. breathy b. modal

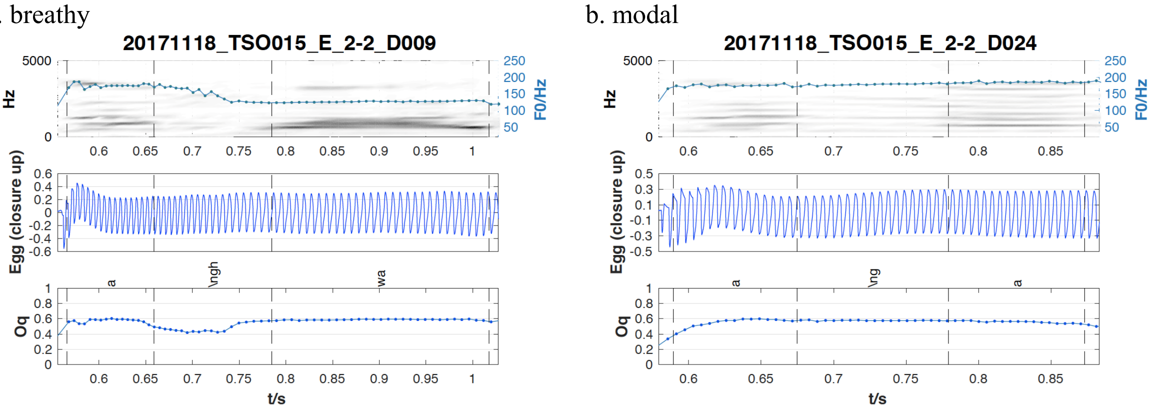

Figure 9: Breathy and modal dorsal nasals in a H tone context by a male speaker

The EGG data in the high tone context from the female speaker (TSO017) shows less reliable Oq patterns, most likely due to thinner vocal folds or incompletely closing vocal folds: figure 10 is for labials, figure 11 is for coronals, and figure 12 is for dorsals. For this female speaker, an observable difference can be found in the duration of the nasal consonants, but not in Oq. While modal nasals are consistently around 100 ms, breathy nasals are longer than modal nasals (ca. 150 ms). This durational difference between breathy and modal voice nasals is not observed in the male speaker who produced most nasals with a duration of around 130 ms.

a. breathy b. modal

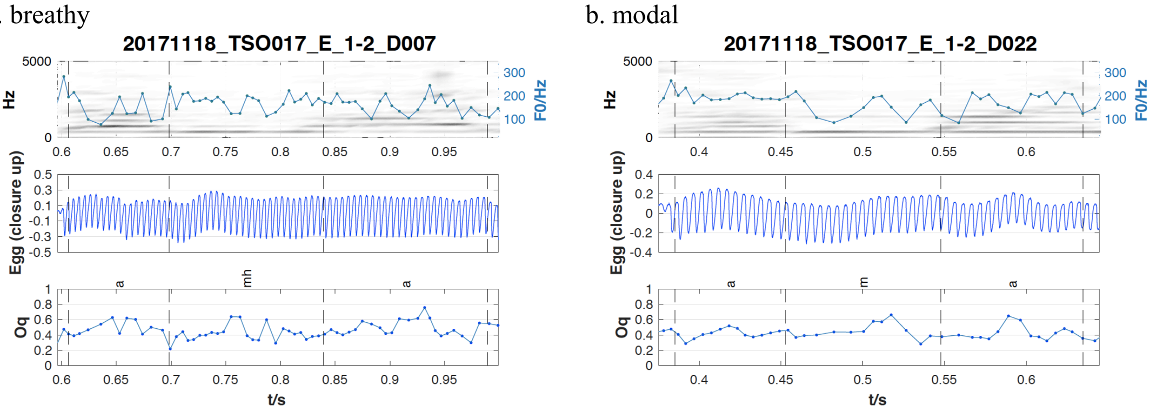

Figure 10: Breathy and modal labial nasals in a H tone context by a female speaker

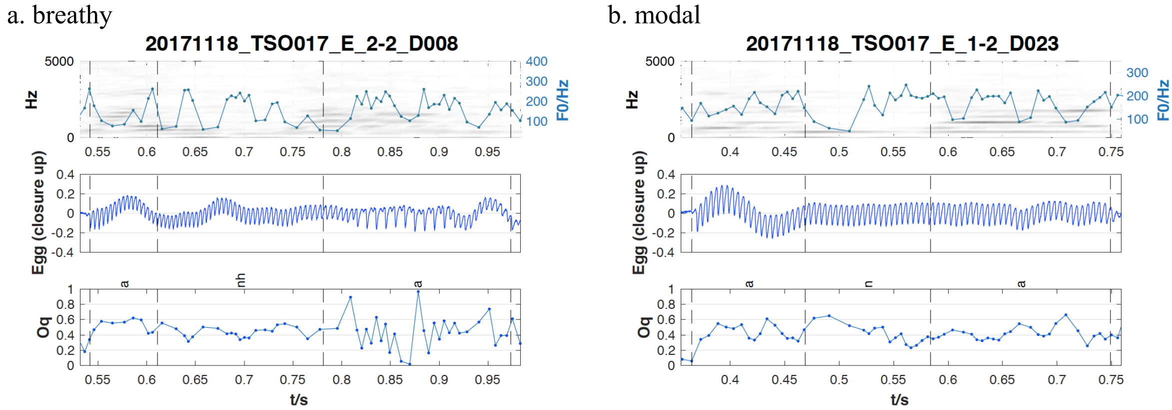

Figure 11: Breathy and modal coronal nasals in a H tone context by a female speaker

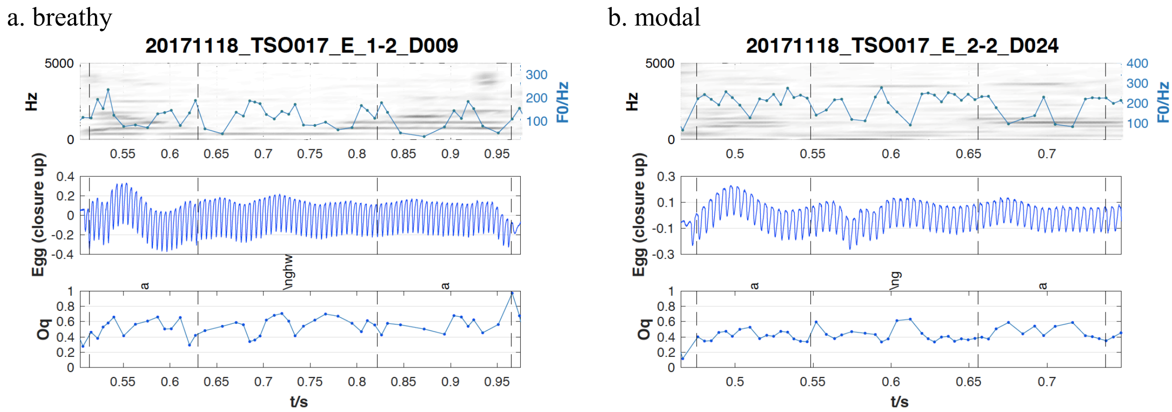

Figure 12: Breathy and modal dorsal nasals in a H tone context by a female speaker

4 Discussion

The results in this study have shown three different patterns concerning the breathy versus modal voice nasals: (a) no distinction (male speaker, toneless context), (b) slight dip in Oq values (male speaker, H tone context), and (c) duration difference (female speaker, H tone context). It is possible that one or more than one of these cues serve as the provider for the phonation contrast, but this needs to be investigated further using studies that test the perception of the contrast.

An interesting observation in the current study lies in the durational difference: breathy nasals were longer than modal nasals in the female speaker's speech. Since breathiness is an indication of flapping movement, it is not unreasonable to think that breathy nasals may require longer time to produce than modal nasals. What is interesting is that the male speaker does not show such durational difference, while the female speaker does. Although it is difficult to generalize based on two speakers, the preliminary findings suggest that Xitsonga speakers may have started a language change where the phonation contrast becomes a length contrast.

The fact that minimal pairs or near-minimal pairs to show phonation contrasts between modal and breathy nasals in Xitsonga do not exist does not exclude the presence of a phonological contrast. Nevertheless, it raises questions as to how exactly the phonation contrast in Xitsonga is realized. Understanding the articulatory nature of breathy nasals is important because breathy nasals also block high tone spreading when a word has a H tone that is preceded by more than one toneless syllable (Lee 2015).

The phonation contrast in Xitsonga is coded in the spelling as breathy nasals are written with an 'h': 'm' is a modal nasal, but 'mh' is a breathy nasal. Xitsonga speakers furthermore show clear judgements as to which word requires this additional 'h' after a nasal. This categorical difference in Xitsonga speakers' grammar, however, may not be realized with a phonation difference, which supports the language change view: the contrast between so-called breathy nasals and so-called modal nasals is maintained through duration rather than types of vocal fold vibration. Such a knowledge is not unusual. The three-way laryngeal contrast in Korean plosives uses different phonetic cues based on speech groups. Voice onset time is a good predictor between the three plosives (unaspirated, aspirated and tensed) in older Korean speakers' speech. On the other hand, younger speakers of Korean categorize the three-way plosive contrast using voice onset time and fundamental frequency (Silva 2006). The three-way contrast in the Korean phonology is maintained but phonetic realizations of these categories have changed. Such a change in Korean could be akin to what has been observed in Xitsonga.

5 Conclusion

This paper reports on results from an articulatory study of the contrasts in Xitsonga nasal phonation between breathy and modal.. It complements an earlier description (Baumbach 1987) and a phonetic study (Traill and Jackson 1988) by measuring the opening and closing of the vocal cords using the Electroglottograph machine while Xitsonga speakers read orthographic representations of the phonation contrast. Although speakers have knowledge of this contrast, EGG and acoustic data do not fully support an analysis that Xitsonga has a salient phonation contrast between modal and breathy nasals. The duration difference found in the female speaker may serve as a phonetic cue for categorizing the two types of nasals, and may possiby indicate ongoing changes in phonetic realizaations of these contrasts. The current study will need to include perception studies in the future so that the role of various acoustic cues utilized by Xitsonga speakers can be discerned.

Acknowledgements

I thank Xitsonga consultants who participated in the study as well as Crous Hlungwani and Dean Mogale Makgopa at University of Venda. I also thank comments from participants of AJL 3 as well as discussions with Shigeto Kawahara, Julián Villegas and Jeremy Perkins. I also thank Michinori Suzuki, Haruka Takeyama and Mana Ashida for annotating Xitsonga data, and Donna Erickson for proofreading the manuscript. This study has partially been funded by a Kakenhi grant 16K02641 and the “Establishment of a Research Network for Exploring the Linguistic Diversity and Linguistic Dynamism in Africa” grant funded by the JSPS's Core-to-Core Program: B. Asia-Africa Science Platforms (2018-2020). All remaining errors are on my own.

References

Baumbach, E. (1987) *Analytical Tsonga*, UNISA.

Laver, John (1980) *The phonetic description of voice quality*. Cambridge University Press.

Lee, S. J. (2008) Consonant-tone interaction in Optimality Theory, Ph.D. Dissertation, Rutgers University.

Lee, S. J. (2015) Cumulative effects in Xitsonga: high-tone spreading and depressor consonants. *Southern African Linguistics and Applied Language Studies* 33 (3), 273-290.

Louw, J. A. (1968) *The intonation of the sentence and its constituent parts in Xhosa and Tsonga*. Pretoria: University of South Africa.

Mathes, T. (2014) Extreme Tonal Depressor Effects in Khoisan: Evidence from Tsua. *Proceedings of the 4th International Symposium on Tonal Aspects of Languages* (TAL-2014) http://www.isca-speech.org/archive

Ohala, J. & M. Ohala (1993) The phonetics of nasal phonology. In: *Phonetics and Phonology* Vol. 5. Academic Press. 225-249.

Silva, David J. (2006) Acoustic Evidence for the Development of Pitch Accent in Standard Korean. *Phonology* 27.2:287-308.

Traill, A. and M. Jackson (1988) Speaker variation and phonation type in Tsonga nasals. *Journal of Phonetics* 16: 385-400.

Villegas, J. (2018) Eggnog v. 0.3. Onkyo Lab, Aizu University.