PROSODIC-STRUCTURAL EFFECTS ON VOICE QUALITY ASSOCIATED WITH KOREAN THREE-WAY STOP CONTRAST

Jiyoung Jang[a,c], Sahyang Kim[b], and Taehong Cho[c]

[a]University of California Santa Barbara; [b]Hongik University; [c]Hanyang Institute for Phonetics and Cognitive Sciences of Language, Hanyang University
jiyoung@ucsb.edu; sahyang@hongik.ac.kr; tcho@hanyang.ac.kr

ABSTRACT

The present study examines the effect of prosodic structure on voice quality associated with the Korean three-way stop contrast (lenis, aspirated, fortis). While the three-way contrast has been known to be signaled primarily by an interaction between VOT and F0, the focus of this study is to explore to what extent the voice quality difference observable in the following vowel would contribute to the three-way stop contrast, and how the realization is conditioned by prominence- vs. boundary-induced prosodic strengthening. Results of twelve young native Korean speakers suggest that the difference in voice quality remains largely significant between the three stop categories, indicating the role of voice quality in the maintenance of the stop contrast. Furthermore, prosodic strengthening shows different effects on the contrast, depending on whether the source comes from prominence or boundary. Some implications will also be discussed in relation to the on-going tonogenetic sound change in Korean.

Keywords: voice quality, breathiness, Korean, three-way stop contrast, prosodic structure

1. INTRODUCTION

Korean has a three-way voiceless stop contrast between aspirated, fortis, and lenis stops [3, 16]. VOT, among other cues, has been regarded as the primary cue which distinguished the three stop categories. In recent studies, however, it has been suggested that Korean is undergoing a tonogenetic sound change, showing that the VOT distinction between the aspirated and lenis stop is merging and that the contrast is now distinguished by F0 (higher F0 for aspirated and lower F0 for the lenis) in younger Koreans' speech [2, 16].

Traditionally, however, the three-way stop contrast in Korean has also been known to be phonetically characterized by the difference in voice quality of the following vowel, such that the breathiness of the vowel is greatest after the lenis, intermediate after the aspirated, and least after the fortis [3]. The purpose of the present study is to examine these voice quality differences associated with the three-way contrastive stops, and to explore to what extent the voice quality difference contributes to the three-way stop contrast and how the effect may be conditioned by two prosodic-structural factors: focus-induced prominence and prosodic boundary, both of which are known to strengthen segmental realization in linguistically meaningful ways [4, 17]

Regarding the effect of prosodic strengthening on the voice quality of the three stops, it is possible to hypothesize that the strengthening would be phonetically manifested by increased glottalization across the board as sounds are known to be glottalized (or become creakier) in prosodic strengthening environments [5, 9, 12, 20]. Previous studies on English [9, 12, 20] have shown that the degree of glottalization in word-initial vowels is greater under accent and/or in domain-initial positions. Similarly, Cho et al. [5] showed that both prominence and prosodic boundary increase the degree of glottalization of initial vowels in South Kyungsang Korean. These results are consistent with a view that prosodic strengthening involves an increase in articulatory force which applies to both laryngeal and supralaryngeal articulation [10]. The prosodic strengthening effect may therefore increase the laryngeal muscular tension, which in turn would augment the degree of glottalization. If the three-way contrastive stops are produced with an increase in laryngeal muscular tension, the breathiness of the following vowel across the three-way stop categories would be reduced.

Previous studies of prosodic strengthening, however, have also indicated that prosodic strengthening is not a mere low-level phonetic effect, but it makes reference to the phonological system of a given language, often enhancing the phonological contrast [6, 17]. This leads to an alternative hypothesis that prosodic strengthening would increase the breathiness of the vowel only after the lenis stop while the creakiness of the vowel is reinforced after the fortis stop, enhancing the three-way stop contrast. This alternative hypothesis may not be borne out, however, if VOT and F0 are the only primary phonetic cues of phonological contrast in Korean, as has been argued in conjunction of a recent development of the tonogenetic sound change among young speakers [2, 16].

To explore these possibilities, the present study investigates whether the three-way phonological contrast of word-initial stops manifest itself in the voice quality of the following vowel in (Seoul) Korean in prosodic strengthening environments. The amplitude differences between the first and second harmonics

(H1*-H2*) and between the first harmonic and F1 (H1*-A1*), were taken as acoustic indexes of the voice quality with higher value indicating greater breathiness in the vowel [11]. Given that the degree of breathiness is often inversely correlated with the degree of creakiness [3, 13], these measures will also be used to assess where in the breathy-creaky continuum the following vowel of each stop category falls.

2. METHOD

2.1. Participants, materials, and procedure

Twelve native Korean speakers (7F, 5M; Mean age = 24.4, age range = 21-28) participated in the study. They were undergraduate students at the time of the recording. Four triplets of Korean monosyllabic CVC words were used as in Table 1. Most of them were nonce words created to meet the following criteria for a larger corpus: The onset of the word was one of the three-way contrastive stops: lenis (/p, t/), aspirated (/p^h, t^h/), or fortis (/p*, t*/). The following vowel was controlled as /a/ and the coda consonant was either /k, t, p/.

Table 1: Target words used for the present study.

Lenis	Aspirated	Fortis
박 /pak/	팍 /p^hak/	빡 /p*ak/
밧 /pat/	팟 /p^hat/	빳 /p*at/
답 /tap/	탑 /t^hap/	땁 /t*ap/
닷 /tat/	탓 /t^hat/	땃 /t*at/

The target words were placed in carrier sentences with different prosodic renditions (Table 2). Prompt sentences (A in Table 2) were used to help forming a context of the mini-dialogue in which Speakers A and B were playing some kind of board game with cards. As for the boundary conditions (i.e., to test domain-initial strengthening effects), the target word was placed either in phrase-initial position (IP-initial) or in phrase-medial position (IP-medial). For prominence, it either received focus by contrasting the onset consonant with /m/ (the focused condition) or unfocused by placing a focused element in another location in the sentence (the unfocused condition).

Table 2: Example sentences with the target word *pak* in different prosodic context. Focused words are in bold.

conditions	example sentences
IP-i, Foc	A: [ipʌn tanʌnɯn **mak**satʃʲin twienonni]? "This time, do I place the word (card) to the right of the picture of **mak**?" B: [ani]. IP [**pak**satʃʲin twi]. IP [twɛssʌ]? "No. To the right of the picture of **pak**. Got it?"
IP-i, Unfoc	A: [ipʌn tanʌnɯn paksatʃʲin **ap^{h}e**nonni]? "This time, do I place the word (card) to the **left** of the picture of pak?" B: [ani]. IP [paksatʃʲin **twi**]. IP [twɛssʌ]? "No. To the **right** to the picture of pak. Got it?"
IP-m, Foc	A: [ipʌn tanʌnɯn ap*a **mak**satʃʲin twienonni]? "This time, do I place the word (card) to the right of dad's picture of **mak**?" B: [ani]. IP [a*pa **pak**satʃʲin twi]. IP [twɛssʌ]? "No. To the right of dad's picture of **pak**. Got it?"
IP-m, Unfoc	A: [ipʌn tanʌnɯn ap*a pak.satʃʲin **ap^{h}e**nonni]? "This time, do I place the word (card) to the **left** of dad's picture of pak?" B: [ani]. IP [a*pa paksatʃʲin **twi**]. IP [twɛssʌ]? "No. To the **right** of dad's picture of pak. Got it?"

Speakers were asked to produce the test sentences in response to the prompt questions. Instead of the full written texts of carrier sentences, some visual clues for the carrier sentences were provided on a computer screen. For example, the screen showed two cards on which a monosyllabic test word was written on each of them in a contrastive way (e.g., *pak* vs. *mak*). The target word (e.g., *pak*) was marked with "O" and its contrasting word (e.g., mak) with "X". The pre-recorded voice was played through the loudspeaker, asking the speaker whether the next word to pick would be the contrasting word (with "X"). The speaker, cued by an "O" mark on the correct (target) word on the screen, was instructed to correct it by saying that the other one should be picked, thus making (corrective) focus on the target word. Given that the carrier sentences were simple, participants were able to produce the intended sentences in response after having received an about 10-minute training session. The prompt sentences were recorded prior to the experiment by two native speakers (1F, 1M). Acoustic data were collected in a soundproof booth using a Tascam HC-P2 digital recorder and a SHURE KSN44 condenser microphone at a sampling rate of 44kHz. In total, 2304 tokens were collected (12 target words x 2 boundary types x 2 focus types x 4 repetitions x 12 speakers), and 2037 tokens were used for further analysis, discarding tokens with unintended prosodic rendition.

2.2. Measurements and statistical analyses

H1*-H2* and H1*-A1* were measured as indexes of the degree of breathiness (and creakiness), obtained by VoiceSauce [21, 22]. Note that * here indicates corrected measures for the effect of formant frequencies [14, 15]. The values were obtained at the 25% and 50% points of the vowel. Repeated Measures ANOVAs were carried out with Stop (aspirated, fortis, lenis), Focus (focused, unfocused), and Boundary (IP-initial, IP-medial) as within-subject factors, and pair-wise t-tests were carried out to examine where the interaction came from.

3. RESULTS

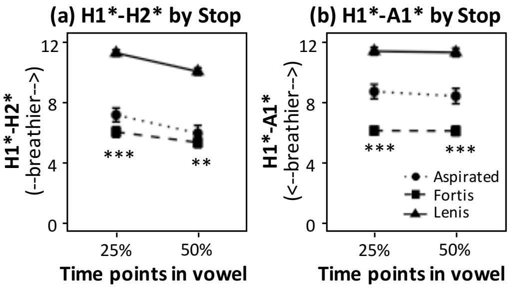
For the purpose of the present study, we will only report results that are directly related to the research questions (i.e., main effects of Stop and its interaction with Focus and Boundary).

3.1. H1*-H2*

There was a significant main effect of Stop on H1*-H2* at both 25% and 50% points in the vowel (25%,

$F[2,22]=27.65$, $p<.001$; 50%, $F[2,22]= 13.71$, $p<.01$), suggesting that the difference on vowel quality among the three-way stop categories remains significant. Vowels were statistically most breathy (greatest H1*-H2*) after the lenis stop and least breathy (smallest H1*-H2*) after the fortis stop. As shown in Fig.1a, H1*-H2* was intermediate for the aspirated stop at the 25% point, showing a three-way contrast, but it did not differ from that of the fortis stop at the 50% point.

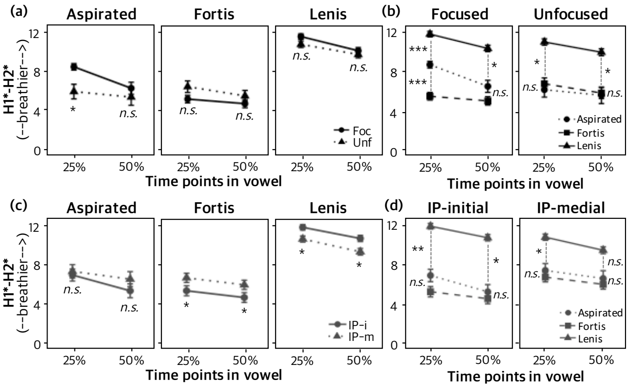
Figure 1: Effect of Stop on (a) H1*-H2* and (b) H1*-A1*. ('**' refers to $p<.01$ and '***' to $p<.001$).



Effect of prominence. There was a significant interaction between Stop and Focus on H1*-H2* at the 25% point ($F[2,22]=7.19$, $p<.01$). The interaction was due to the fact that the focus effect was significant only at the 25% point after the aspirated stop. As shown in Fig. 2a, H1*-H2* value for the aspirated stop was smaller in the focused than in the unfocused condition (at 25%), indicating that the breathiness of the aspirated stop was increased under focus. From a different perspective, the interaction at the 25% point was also in part due to the difference in the effect size of Stop in the focused vs. unfocused conditions: the difference between the stops was larger in the focused condition (eta^2=.75; mean diff., aspirated vs. lenis=-2, aspirated vs. fortis=4, lenis vs. fortis=6) than in the unfocused condition (eta^2=.53, aspirated vs. lenis=-5, aspirated vs. fortis=0, lenis vs. fortis=4) (Fig. 2b).

Effect of boundary (domain-initial). There was an interaction between Stop and Boundary at both 25% and 50% time points (25%, $F[2,22]=5.07$, $p<.05$; 50%, $F[2,22]=4.81$, $p<.05$). As shown in Fig. 2c, the boundary effect was significant for the fortis and the lenis stops, but not for the aspirated stop. Interestingly, the presence of a larger boundary had an opposite effect for the fortis vs. the lenis stop: IP-initially, the vowels (compared to the IP-medial ones) showed smaller H1*-H2* (less breathy) after the fortis stop, but larger H1*-H2* (more breathy) after the lenis stop. The interaction was also attributable to the fact that the effect of Stop was greater in the IP-initial position (25%, eta^2=.685; mean diff., aspirated vs. lenis=-5, aspirated vs. fortis=2, lenis vs. fortis=6; 50%, eta^2, =.58; mean diff., aspirated vs. lenis=-5, aspirated vs. fortis=1, lenis vs. fortis=6) than in the IP-medial position (25%, eta^2 =.517; mean diff., aspirated vs. lenis=-3, aspirated vs. fortis=1, lenis vs. fortis=4; 50%, eta^2=.365; mean diff., aspirated vs. lenis=-2, aspirated vs. fortis=0, lenis vs. fortis=3).

Figure 2: Effects of prosodic factors on H1*-H2*. The Stop x Focus interaction is illustrated in (a) by the stop category and in (b) by the focus condition. The Stop x Boundary interaction is illustrated in (c) by the stop category and in (d) by the boundary condition. Note that the difference between the lenis and the fortis was significant in all cases. ('*'=$p<.05$).



3.2. H1*-A1*

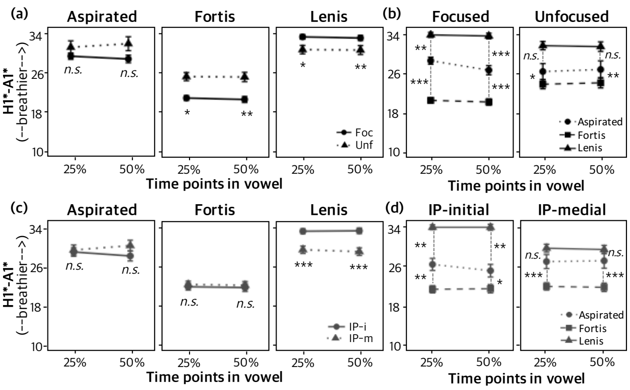
A significant main effect of Stop was found at both 25% and 50% points in the vowel (25%, $F[2,22]=85.89$, $p<.001$; 50%, $F[2,22] =67.87$, $p<.001$) showing a three-way distinction among the stops. As can be seen in Fig.1b, vowels were statistically most breathy (greatest H1*-A1*) after the lenis stop and least breathy (smallest H1*-A1*) after the lenis stops. H1*-A1* was intermediate for aspirated stops, showing a three-way stop contrast in H1*-A1*.

Effect of prominence. There was a significant Stop x Focus interaction at both measurement points (25%, $F[2,22]=6.76$, $p<.01$; 50%, $F[2,22]=5.66$, $p<.05$). The interaction was due to the fact that the focus effect was significant in the vowel after the fortis and the lenis stop, but not after the aspirated stop (Fig. 3a). The effect showed an opposite direction for the fortis vs. the lenis stop, with the focus effect of decreasing H1*-A1* for the fortis stop (thus being less breathy/creakier under focus), but of increasing H1*-A1* for the lenis stops (being more breathy under focus). Another attribute to the interaction was the different effect sizes of Stop in the focused vs. unfocused conditions. As shown in Fig. 3b, the Stop effect was larger in the focused condition (25%, eta^2=.92; mean diff., aspirated vs. lenis=-5, aspirated vs. fortis=9, lenis vs. fortis=14; 50%, eta^2=.91; mean diff., aspirated vs. lenis=-7, aspirated vs. fortis=7, lenis vs. fortis=14) compared to unfocused condition (25%, eta^2=.58; mean diff., aspirated vs. lenis=-5, aspirated vs. fortis=3, lenis vs. fortis=8; 50%, eta^2=.57; mean diff., aspirated vs. lenis=-5, aspirated vs. fortis=3, lenis vs. fortis=8).

Effect of boundary (domain-initial). There was a Stop x Boundary interaction at both 25% and 50% time points (25%, $F[2,22]=6.91$, $p<.05$; 50%, $F[2,22]=9.87$,

p<.01). The interaction was due to the boundary effect being only significant after the lenis stop. H1*-A1* after the lenis stop was larger (breathier) in the IP-initial than in the IP-medial position (Fig. 3c). Moreover, the effect of Stop was larger in the IP-initial position (25%, eta^2=.8; mean diff., aspirated vs. lenis=-8, aspirated vs. fortis=6, lenis vs. fortis=14; 50%, eta^2=.81; mean diff., aspirated vs. lenis=-9, aspirated vs. fortis=5, lenis vs. fortis=14) than in the IP-medial position (25%, eta^2=.73; mean diff., aspirated vs. lenis=-2, aspirated vs. fortis=6, lenis vs. fortis=8; 50%, eta^2=.74; mean diff., aspirated vs. lenis=-3, aspirated vs. fortis=6, lenis vs. fortis=8) (Fig. 3d), showing clear stop distinction in the domain-initial position.

Figure 3: Effects of prosodic factors on H1*-A1*. The Stop x Focus interaction is illustrated in (a) by the stop category and in (b) by the focus condition. The Stop x Boundary interaction is illustrated in (c) by the stop category and in (d) by the boundary condition. Note that the difference between the lenis and the fortis was significant in all cases. (‘*’=p<.05;‘**’=p<.01;‘***’=p<.001).



4. DISCUSSION

One of the basic findings of the present study is that the difference in voice quality of the following vowel, as indicated by breathiness measures (H1*-H2* and H1*-A1*), makes a three-way phonetic distinction among the (Seoul) Korean stop series produced by young speakers. The amount of breathiness is largest for the lenis stop, intermediate for the aspirated stop, and smallest for the fortis stop. The three-way laryngeal distinction is consistent with what was reported 16 years ago [3]. This indicates that the voice quality difference has continued to underlie the three-way stop contrast. This finding is interesting, given the purported tonogenetic sound change which has arguably reduced the three-way phonetic distinction to two primary phonetic cues F0 and VOT (e.g., [2, 16]) in such a way that, for example, the difference between the lenis and the aspirated stops is signalled primarily by F0 with no difference in voice quality. The results of the present study, however, demonstrate that the Korean stop contrast is still characterized by the laryngeal contrast at least at the phonetic level.

Another important finding of the present study is that the three-way distinction in voice quality is further conditioned by prosodic strengthening factors: focus-induced prominence and boundary. The prosodic strengthening effects allow us to understand the phonological role of the phonetic distinction in voice quality. For example, de Jong and colleague [7, 8] suggest that one way to assess the role of phonetic features in making phonological contrast may be to examine whether the phonetic feature participates in enhancing phonological contrast under focus-induced prominence. The results of the present study showed that the three-way stop contrast is indeed enhanced under focus, with the stops being substantially dispersed along the breathy-creaky phonetic continuum. The dispersion effect was observed in both H1*-H2* and H1*-A1*.

The voice quality difference as a function of stop categories has also been found to be influenced by prosodic boundary (IP-initial vs. IP-medial). As was seen in the results section, the exact detail of how the three-way voice quality distinction was modulated by boundary was somewhat different from the case of the prominence-driven strengthening effect. Boundary effect was mainly observed in H1*-H2*, whereas prominence effect was mainly observed in H1*-A1*, showing prosodic strengthening may have different effects as a function of its source: prominence vs. boundary (Fig. 2a,3a vs. Fig. 2c,3c). But as was the case with the prominence effect, the boundary-related strengthening effect also induces an enhancement of the three-way stop contrast, showing some degree of augmented dispersion of the stops along the breathy-creaky continuum.

The converging enhancement pattern under prosodic strengthening is that vowels become more creaky (less breathy) after the fortis stop, but less creaky (more breathy) after the lenis stop, contributing to the enhancement of the phonological contrast. Interestingly, the voice quality associated with the aspirated stop falls somewhere in between, which may be understood as an effort to retain the contrast by maintaining its intermediate position.

The results taken together imply that variation in the voice quality difference as a function of prosodic strengthening is not a mere low-level phonetic effect that would otherwise have applied to all three stops in a collective way, but is an outcome of the phonetic-prosody interface in reference to the phonological contrast in the language. The results also suggest that understanding the nature of laryngeal (voicing) contrast that occurs in Korean as well as in other languages requires multi-dimensional approaches to explore the phonetic realization of both the primary and other non-primary phonetic features [e.g., 1, 18, 19]. It remains to be seen to what extent the voice quality difference is exploited by the listeners and how the voice quality cues may interact with F0 and VOT [cf. 19].

6. REFERENCES

[1] Al-Tamimi, J., & Khattab, G. 2018. Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics*, 71, 306-325.

[2] Bang, H.-Y, Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T.-J. 2018. The emergence, progress, and impact of sound change in progress in Seoul Korean: implications for mechanisms of tonogenesis. *Journal of Phonetics*, 66, 120-144.

[3] Cho, T., Jun, S.-A., & Ladefoged, P. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, 30, 193-228.

[4] Cho, T., Kim, D., & Kim, S. 2017. Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics*, 64, 71-89.

[5] Cho, T., Kim, D., & Kim, S. 2018. Prosodic strengthening in reference to the lexical pitch accent system in South Kyungsang Korean. *The Linguistic Review*.

[6] Cho, T., & McQueen, J. M. 2005. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33(2), 121-157.

[7] de Jong, K. 2004. Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. *Journal of Phonetics*, 32(4), 493-516.

[8] de Jong, K., & Zawaydeh, B. 2002. Comparing stress, lexical focus, and segmental focus: patterns of variation in Arabic vowel duration. *Journal of Phonetics*, 30(1), 53-75.

[9] Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24(4). 423–444.

[10] Fougeron, C. 1999. Prosodically conditioned articulatory variations: a review. *UCLA Working Papers in Phonetics*, 97, 1–74.

[11] Garellek, M. (2013). *Production and perception of glottal stops* (Doctoral dissertation, UCLA).

[12] Garellek, M. 2014. Voice quality strengthening and glottalization. *Journal of Phonetics*, 45, 106-113.

[13] Gordon, M., & Ladefoged, P. 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics*, 29(4), 383-406.

[14] Iseli, M. & Alwan, A. 2004. An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation. *Proceedings of ICASSP*, 669-672, Montreal, Canada.

[15] Iseli, M., Shue, Y. L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America*, 121(4), 2283-2295.

[16] Kang, Y. 2014. Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, 45, 76-90.

[17] Kim, S., Kim, J., & Cho, T. 2018. Prosodic-structural modulation of stop voicing contrast along the VOT continuum in trochaic and iambic words in American English. *Journal of Phonetics*, 71, 65-80.

[18] Kirby, J. P. 2018. Onset pitch perturbations and the cross-linguistic implementation of voicing: Evidence from tonal and non-tonal languages. *Journal of Phonetics*, 71, 326-354.

[19] Kong, E. J., Beckman, M. E., & Edwards, J. 2011. Why are Korean tense stops acquired so early?: The role of acoustic properties. *Journal of Phonetics*, 39(2), 196-211.

[20] Pierrehumbert, J. & Talkin, D. 1992. Lenition of /h/ and glottal stop. *In Gerard J. Docherty & D. Robert Ladd (eds.), Papers in laboratory phonology II: Gesture, segment, prosody*, 90–117. Cambridge, New York & Victoria: Cambridge University Press.

[21] Shue, Yen-Liang 2010. *The voice source in speech production: Data, analysis and models*. Los Angeles: University of California. http://www.phonetics.ucla.edu/voiceproject/Publications/shue_dissertation.pdf

[22] Shue, Y.-L., Keating, P., Vicenik, C. & Yu, K. 2011. Voicesauce: A program for voice analysis. In Wai-Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, 1946–1849. Hong Kong.