Using psychoacoustic roughness to measure creakiness in Burmese

Jeremy Perkins, University of Aizu
Seunghun J. Lee, International Christian University & University of Venda
Julián Villegas, University of Aizu
Kosei Otsuka, Osaka University

ICPP 2018
Oct. 27, 2018
NINJAL, Tachikawa, Japan

Introduction & Background

- Decreased spectral tilt is commonly used to identify creaky phonation in phonetic studies (i.e. Blankenship, 2002; Kreiman et al., 2007; DiCanio, 2009; DiCanio, 2012; Kuang 2013; Keating et al., 2015).
  - The diff. in amplitude between harmonics of F0 and/or between F1, F2 and F3 correlates with OQ.
- Psychoacoustic roughness is more suitable in identifying phonologically contrastive creakiness.
  - Roughness relates to properties of perception.
- Gruber (2011) measured OQ via EGG, spectral tilt and oral airflow, targeting tone contrasts in Burmese.
  - In isolation and phrase-finally, creaky & checked tones were distinguished from low and high tone.
  - Sentence-medially, no distinction was found.
- Recent findings in Zhuang (Perkins et al. 2017) and White Hmong (Villegas et al. 2017) have shown roughness can successfully identify creaky phonation.
- Roughness is compared with spectral tilt in the ability to identify creaky tones in Burmese.

Methods

A. Participants & Stimuli

- Recordings from 12 native Burmese speakers (6 males, 6 females) in Yangon.
- 78 monosyllabic words as stimuli in isolation and frame sentences (L __ L) as in Gruber (2011).
- 5 rep's x 78 words x 2 contexts = 780 tokens
  - Stimuli were balanced for:
    - Tone (low, high, creaky, checked (glottal coda))
    - Coda type (open vs. nasal)
    - Vowel quality ([i], [u], [a]).
  - Onsets varied (mostly alveolar; never labial).
- Words were shown to participants in Burmese script via a slide-show on a laptop.

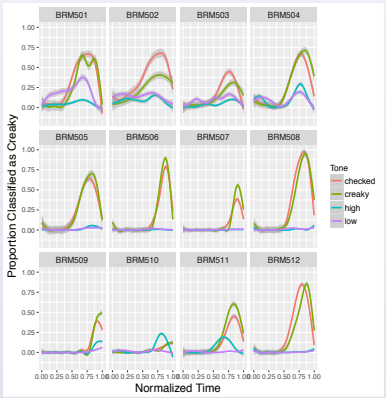
B. Analysis

- Normalized spectral tilt (H1*–H2*, H1*–A1*, **H1*–A2*,** H1–A3*)
  - Using VoiceSauce (Shue et al. 2011) .
- Roughness (binary classification)
  - Using a Matlab script based on Daniel & Weber's (1997) optimization of von Aures' (1995) model.
- % of creakiness measured using COVAREP algorithm.
  - Implementation in Matlab (Degottex et al. 2014, Drugman et al. 2014).
  - Uses a combination of acoustic features.

Results

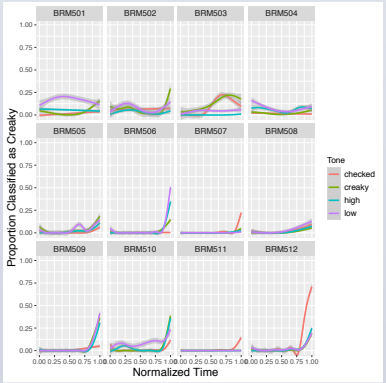
- The roughness algorithm confirmed two major findings in Gruber (2011):
  - Burmese creaky and checked tones have late creakiness in words in isolation.

1) Binary Roughness Classification (in isolation)



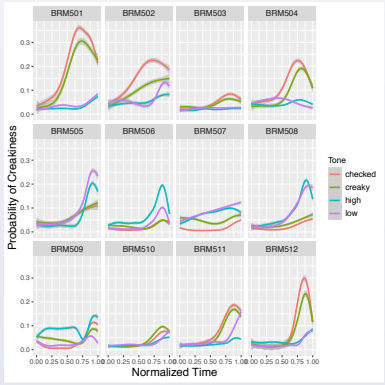
  - For words in sentence-medial context, there was no evidence of creakiness.

2) Binary Roughness Classification (in frame sentences)



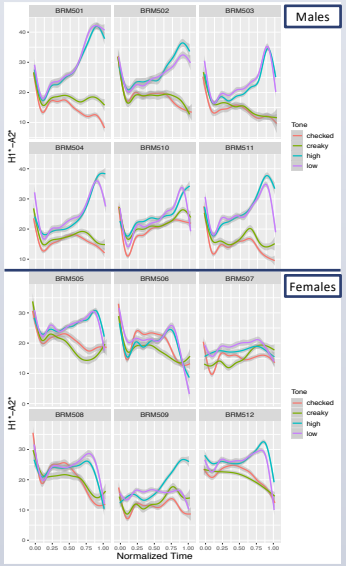
- The COVAREP algorithm did not reliably distinguish creaky from modal tones in most speakers.

3) COVAREP Prob. of Creakiness (in isolation)



- Spectral tilt (H1* – A2*) did not reliably distinguish creaky from modal tones in all speakers.
  - More reliable for males than females:

4) H1* – A2* (in isolation) by gender



Conclusion

- Roughness and not spectral tilt reliably distinguished creaky from non-creaky tones in Burmese.
- Results here match Gruber's (2011) findings:
  - Creakiness is found late in syllables in checked & creaky tones.
  - Creakiness is only found in words read in isolation, and not in frame sentences.

References

Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, *30*(2), 163-191.

Daniel, P. and Weber, R. 1997. Psychoacoustical roughness: implementation of an optimal model. *Acta Acustica United with Acustica*, 83: 113-123.

DiCanio, C. T. 2009. The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association*, *39*(2), 162-188.

DiCanio, C. T. 2012. Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics*, *40*(1), 162-176.

Degottex, G., Kane, J., Drugman, T., Raitio, T. and Scherer, S. 2014. COVAREP – A collaborative voice analysis repository for speech technologies. In *IEEE Int. Conf. on Acoustics, Speech and Signal Processing, ICASSP*, pp. 960-964.

Drugman, T., Kane, J., and Gobl, C. 2014. Data-driven detection and analysis of the patterns of creaky voice. *Computer Speech & Language*, 28(5): 1233-1253.

Gruber, J. 2011. *An articulatory, acoustic, and auditory study of Burmese tone*, Ph.D. dissertation, Georgetown University.

Keating, P, M. Garellek and J. Kreiman. 2015. Acoustic properties of different kinds of creaky voice, *Proc. of the 18th ICPhS*, Glasgow, Scotland.

Kreiman, J., Gerratt, B. R., & Antoñanzas-Barroso, N. 2007. Measures of the glottal source spectrum. *Journal of speech, language, and hearing research*, *50*(3), 595-610.

Kuang, J. 2011. Phonation Contrast in Two Register Contrast Languages and Its Influence on Vowel Quality and Tone. In *ICPhS* (pp. 1146-1149).

Perkins, J., S.J. Lee and J. Villegas. 2017. Psychoacoustic roughness as a measure of creakiness in two dialects of Zhuang. 2017 *Proc. of SICSS 2017*, pp. 111-112.

Shue, Y.-L., P. Keating, C. Vicenik, K. Yu. 2011. VoiceSauce: A program for voice analysis. *Proc. of the ICPhS* XVII, 1846-1849.

Villegas, J., J. Perkins and S.J. Lee. 2017. Psychoacoustic roughness as proxy of creakiness in White Hmong. *Proc. of SICSS 2017*, pp. 74-75.

Von Aures, W. 1985. Berechnungsverfahren für den sensorischen Wohlklang beliebiger Schallsignale (A Model for Calculating the Sensory Euphony of Various Sounds). Acustica, 59:130-141.

Acknowledgement

This work was supported by JSPS Kakenhi Grant Number 15K16745.

Jeremy Perkins – jperkins@u-aizu.ac.jp
Seunghun J. Lee – seunghunjlee@gmail.com
Julián Villegas – julian@u-aizu.ac.jp
Kosei Otsuka – mrkosei@gmail.com