22nd Himalayan Languages Symposium
June 8-10, 2016
Indian Institute of Technology Guwahati

The roles of phonation and f0 in Wuming Zhuang tone

Jeremy Perkins (University of Aizu, Japan)
jperkins@u-aizu.ac.jp
Seunghun J. Lee (International Christian University, Tokyo, Japan)
Julián Villegas (University of Aizu, Japan)

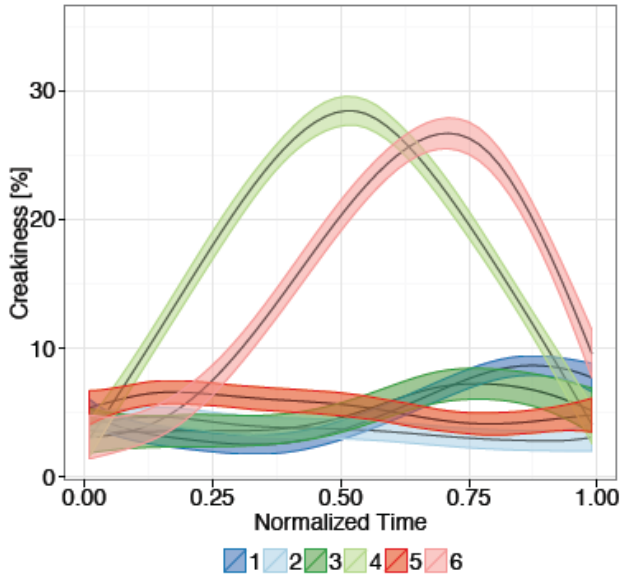
Acoustic correlates of tone

- Fundamental Frequency (F0) (cf. Yip 2002)
  - peak delay
    - Mandarin Chinese (Xu 1998, 1999), Chichewa (Kim 1998, Myers 1999) and Yoruba (Akinlabi and Liberman 1995)
  - downstep and declination

- Phonation and tone (cf. Silverman 1997)
  - laryngeally complex languages
    - Otomanguean (e.g. Jalapa Mazatec, Garellek & Keating 2011; Triqui DiCanio 2008)
    - Nilotic (e.g. Dinka, Andersen 1993)

- Other acoustic correlates
  - duration

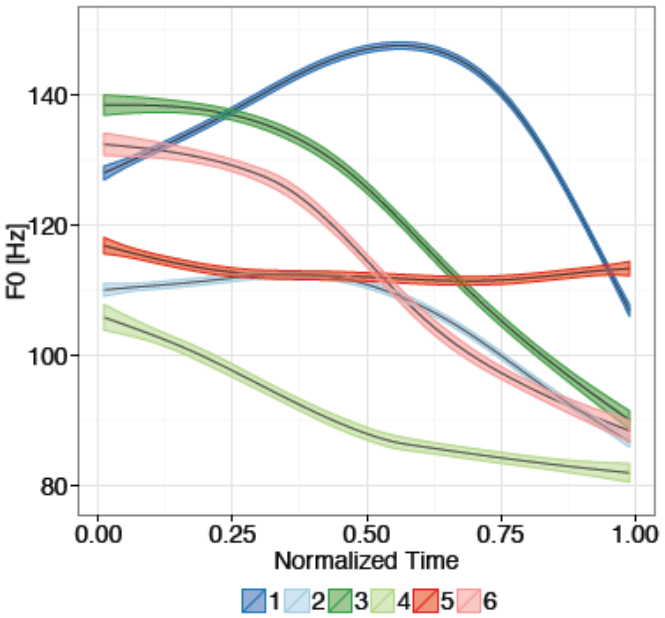
Creakiness Detection Algorithm



- a composite method used to measure creakiness
  - creakiness is estimated every 10 ms using the method in Kane et al. (2013) and Drugman et al. (2014).
  - the algorithm effectively determines the odds of a frame being creaky based on a combination of acoustic features
    - the difference between the first two harmonics (H2-H1)
    - F0 contour
    - residual peak prominence
    - a group of features used by Ishi et al (2008)
      - power peak parameters, inter-pulse similarity, intra-frame periodicity
    - To minimize false positives, three measures are included
      - normalized signal energy, number of zero-crossings, variance in the very short-term power contour
  - Information theoretic methodology was applied to assess how well a set of acoustic features correlated with actual creakiness judgments of recordings from various databases.

SSANOVA

- a smoothing cubic spline ANOVA model (Gu 2014)
  - **F0** and **creakiness** are explained by the factors *Tone* and *normT*, and their interaction
  - smoothing parameters are selected by a generalized cross-validation method using the default parameter values (e.g. $\alpha = 1.4$)

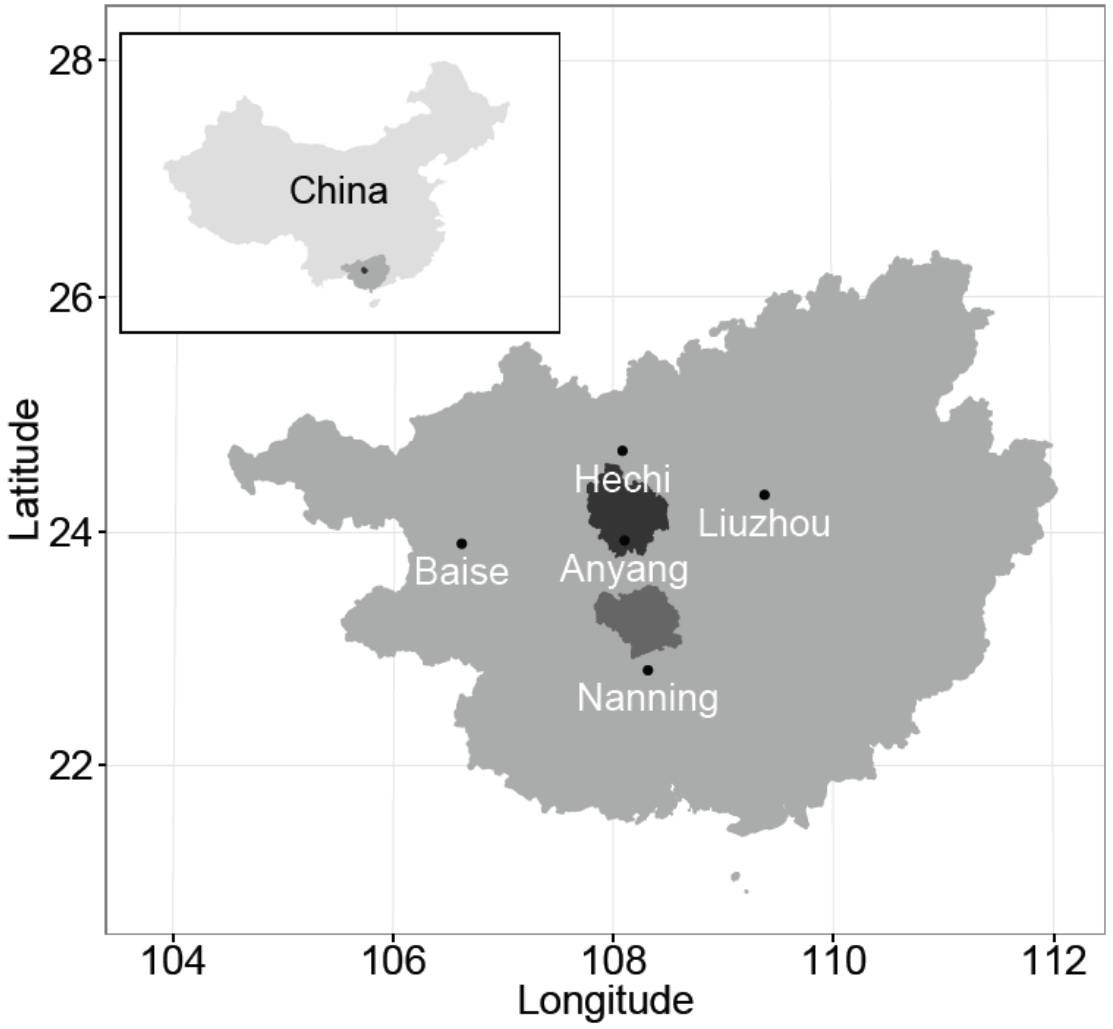


- this method was used in analyses of the lingual and labial articulation of whistled fricatives (Lee-Kim et al. 2014)
- F0 contours and larynx height for Mandarin tones (Moisik et al. 2013)

Outline

- Wuming Zhuang
  - a puzzle in tone categorization
  - the timing of F0 fall
  - the role of creakiness
- Discussion on methodology
  - SSANOVA
  - Creakiness detection algorithm

WUMING ZHUANG



Wuming Zhuang

- Zhuang has the largest number of speakers of the 55 official minority languages in China
  - Zhuang is in the Tai-Kadai family (Thai, Laos, Vietnam, Myanmar & China)
  - The variety spoken in Wuming is considered the standard variety (Wei & Qin, 1980).
  - There is a vast degree of dialectal difference within Zhuang.
  - Many Zhuang dialects are not mutually intelligible.

Wuming Zhuang Tone System (Wei & Qin 1980)

Tone	1	3	5	7 short	7 long
Chao (1930)	24	55	35	55	35
Example	[na]	[na]	[na]	[nap]	[na:p]
Gloss	'thick'	'face'	'arrow'	'to put into'	'to be stuck'
Description	rising	level	rising	level	rising
Tone	**2**	**4**	**6**	**8 short**	**8 long**
Chao	31	42	33	33	33
Example	[na]	[na]	[na]	[nap]	[na:p]
Gloss	'field'	'aunt'	'meat'	'to bind'	'to turn in tax
Description	falling	falling	level	level	level
	unchecked syllables			*checked syllables*	

- Tones and syllable structures
  - Tones 1 to 6 only occur in open syllables or with sonorant codas (*unstopped/unchecked syllables*)
  - Tones 7 & 8 only occur with obstruent codas (*stopped/checked syllables*)
- Assumed tonal register split (following the Chinese tradition)
  - Tones 1, 3, 5, and 7 are upper register
  - Tones 2, 4, 6, and 8 are lower register

Data Collection

- Consultants
  - Three female and one male native speakers of Wuming Zhuang in their 20's were recorded in a sound attenuated booth at Guangxi University in 2015. The university is located in Nanning, Guangxi.
  - At the time of the elicitation, the consultants communicated with their relatives and friends from Wuming in Zhuang. In Nanning, however, the consultants mostly used a Guangxi variety of Putonghua (standard Chinese).

- Procedure
  - Wuming Zhuang words were elicited using a frame sentence, presented in Chinese characters (but read in Zhuang).
    - 我正在读__这个词 "I am reading this word ____ now"
  - Words in isolation were then elicited from a list of Chinese characters to exclude possible tone sandhi effects.
  - The tones for these words in Wuming Zhuang are reported, facilitating the analysis.

Data Collection
Tone Tokens and Exclusion criteria

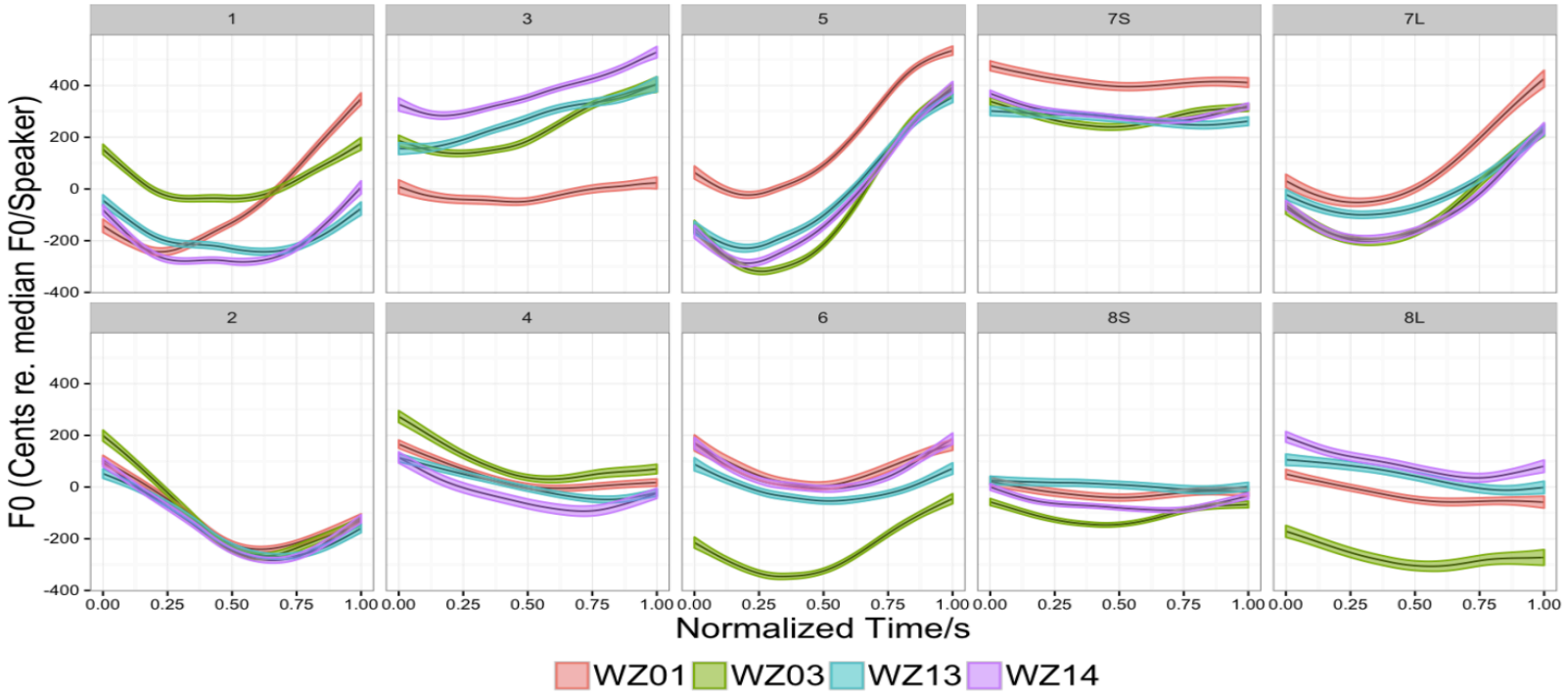
- Monosyllabic words
  - Only monophthongal vowels were included
  - Onset consonants were limited to alveolar, palatal and velar obstruents
  - Nasal codas were elicited but excluded from analysis due to induced creakiness
  - among checked syllables, only alveolar and velar stop codas were included.

- In sum
  - 197 words with 5 repetitions
  - 985 tokens in total (for each speaker)

Acoustic Analysis

- Syllable rhymes were segmented using Praat (Boersma &Weenink 2015)
  - rhyme boundaries are determined using
    - increased intensity
    - appearance and disappearance of the vowel formants
  - A script was run that adjusted all the rhyme boundaries to the nearest zero-crossing.
  - Octave errors were corrected in R by identifying and adjusting points that differed by more than 800 cents from a given speaker's mean.
  - A small number of tokens (66 out of 2627) were excluded because of discontinuities between consecutive F0 measurements.
  - If two consecutive F0 points differed by more than 250 cents, then that token was excluded.

Results – F0



Wei & Qin	1	3	5	7 short	7 long
Tone	24	55	35	55	35
Description	rising	level	rising	level	rising
Tone	2	4	6	8 short	8 long
Chao	31	42	33	33	33
Description	falling	falling	level	level	level
	unchecked syllables			*checked syllables*	

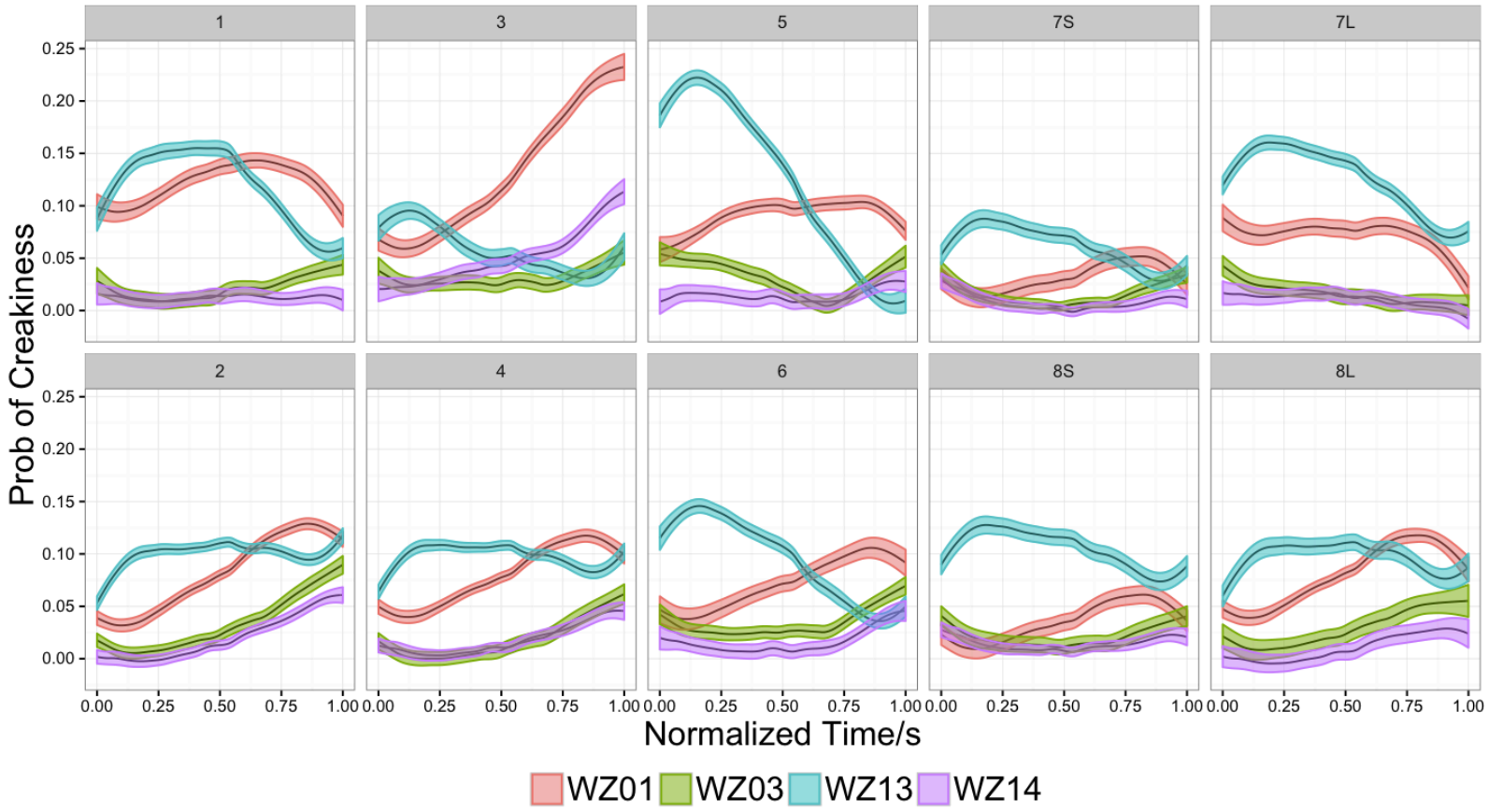
Results – F0

- F0 contour is a good predictor for tone ($r^2 = 0.715$)

- The F0 contours generally agree with Wei & Qin's (1980) account of Wuming Zhuang
  - tones 2 and 4 are classified as falling [31] and [42], consistent with the findings in Figure 1
  - tones 1 and 5 are classified as rising [24] and [35]; our results confirm that F0 rises to a higher point in tone 5 than in tone 1
  - tones 3 and 6, [55] and [33] respectively, have the greatest variation across speakers.
  - checked tones are generally consistent with Wei & Qin's descriptions

- Outlier
  - WZ03 (green) produces tone 6 with a lower F0 and a sharply falling-rising contour.

Results - Creakiness



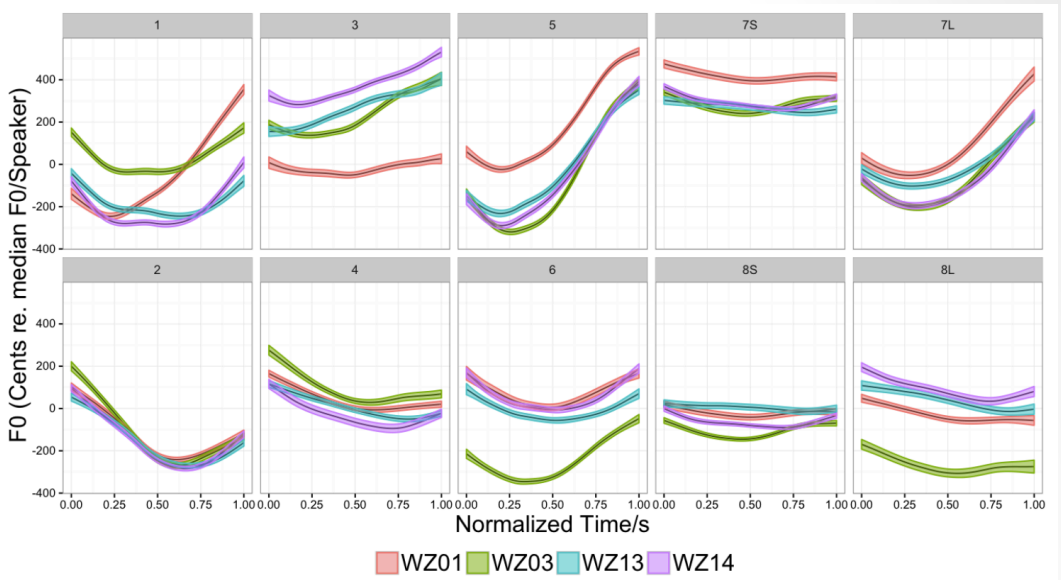
Results - Creakiness

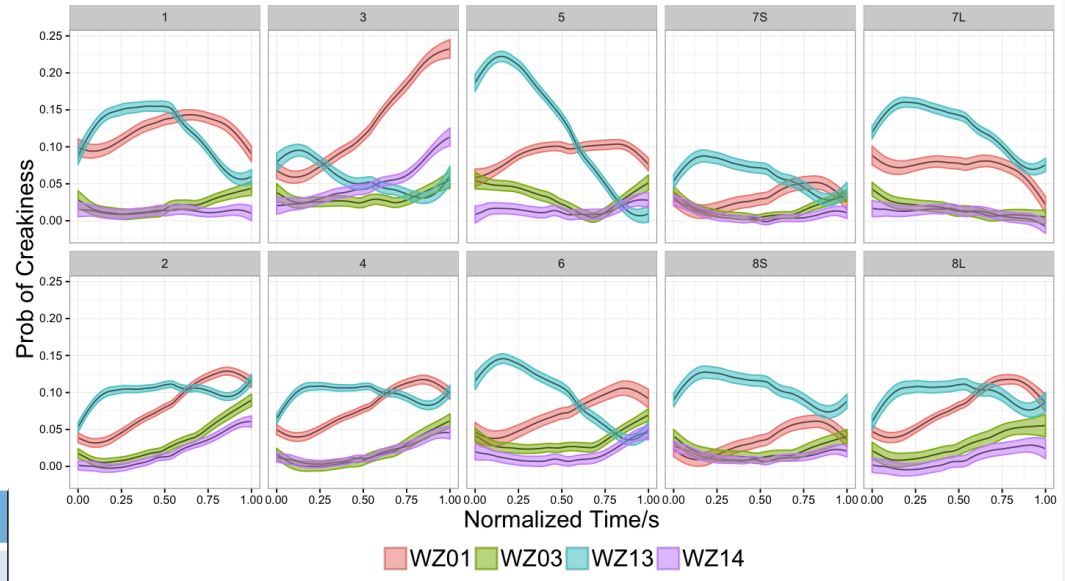
- There is larger variation between speakers in creakiness (than in F0)

- the SS-ANOVA model for creakiness does not fit well (r^2 = 0.332)
  - WZ13 appears to differ from the other speakers in that he is generally creakier than the three female speakers, and has an earlier peak.
  - WZ03 and WZ14 have nearly identical, negligible amount of creakiness for all tones.
  - Only WZ01's tone 3 and WZ13's tone 5 show evidence of significantly increased creakiness

F0

Creakiness



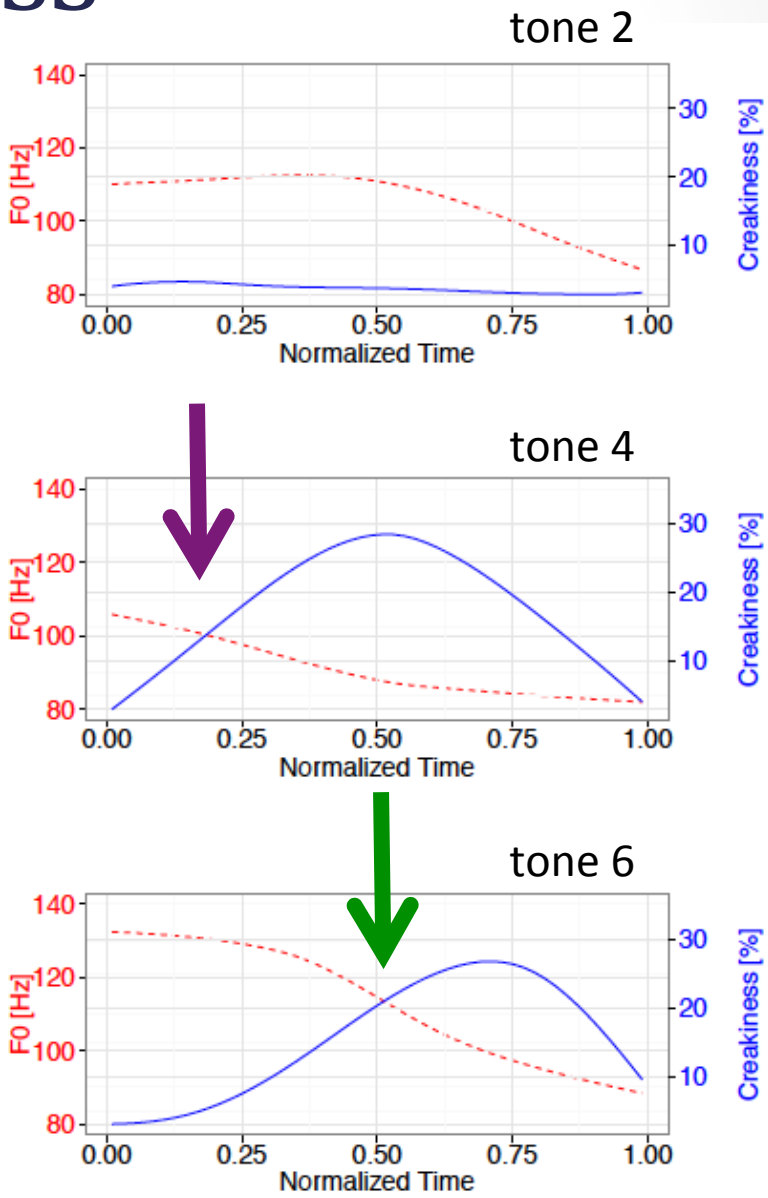


Wei & Qin	1	3	5
Tone	24	55	35
Description	rising	level	rising
Tone	2	4	6
Chao	31	42	33
Description	falling	falling	level

DIALECTAL DIFFERENCES: WUMING VS. DU'AN ZHUANG

F0 and Creakiness

- Overlaid temporal contours of F0 (dashed line) and Creakiness (solid line)
  - tone 2 (top panel)
    - no change in creakiness and F0 drops toward the end of the syllable
  - tone 4 (middle panel)
    - as creakiness increases, F0 falls (purple arrow)
  - tone 6 (bottom panel)
    - as creakiness increases, F0 falls (green arrow, though at a later point compared to tone 4)



Tone 2 – Tone 4 Contrast in Du'an Zhuang

- The difference between tone 2 (31 modal) and tone 4 (31 creaky) lies in the phonation and the timing of the falling of F0.
  - Tone 2 has modal phonation and has a late phonetic fall; it corresponds to [**31**] in Wuming Zhuang.
  - Tone 4 has creaky phonation and has an early phonetic fall; it corresponds to [**42**] in Wuming Zhuang.

- Hypothesis: In Du'an Zhuang, a phonation contrast is in the process of replacing the register contrast based on F0.
  - We are currently in the process of preparing a perception study that assesses the extent of the phonation contrast in Du'an Zhuang and Wuming Zhuang.

Creakiness & F0 in Du'an Zhuang

- The creakiness detection algorithm, coupled with the F0 contour, shows that
  - the timing of when the creakiness increases affects the F0 profile
    - as creakiness increases, F0 drops
  - F0 is dependent on the creakiness profile

- Using measurements such as spectral tilt over the vowel may not provide an accurate picture of the interplay between F0 and creakiness

Conclusion

- Acoustic methods are used to further understand the interplay between phonation and F0, two known acoustic properties that are related to tone.
  - Creakiness detection algorithm
  - SSANOVA
- In Wuming Zhuang, an instrumental investigation of F0 profiles shows that the speakers conform with earlier descriptions in Wei & Qin (1980)
  - Unlike the Du'an Zhuang, however, there is no observable evidence that Wuming Zhuang employs phonation in distinguishing tonal categories.

References

- Akinlabi, A., and M. Liberman (1995) On the phonetic interpretation of the Yoruba tonal system. Proceedings of the International Congress of Phonetic Sciences 1995. Stockholm, Sweden, pp. 42–45.
- Andersen, Torben (1993) Vowel quality alternation in Dinka verb inflection. Phonology 10, 1–42.
- Boersma, Paul (2001) Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345
- Chao, Yuen-Ren (1930) A system of tone letters. *Le Maitre Phonetique* 45, 24-27.
- Degottex, Gilles, Kane, John, Drugman, Thomas, Raitio, Tuomo and Scherer, Stefan (2014) COVAREP – A collaborative voice analysis repository for speech technologies. In *IEEE Int. Conf. on Acoustics, Speech and Signal Processing, ICASSP*, 960-964.
- DiCanio, Christian T. (2008) *The Phonetics and Phonology of San Martín Itunyoso Trique*. PhD dissertation, UC Berkeley.
- Drugman, Thomas, Kane, John and Gobl, Christer (2014) Data-driven detection and analysis of the patterns of creaky voice. *Computer Speech & Language*, 28(5):1233-1253.
- Garellek, Mark & Patricia Keating (2011) The acoustic consequences of phonation and tone interactions in Jalapa Mazatec, Journal of the International Phonetic Association 41(2): 185-205.
- Gu, Chong (2014) Smoothing Spline Anova Models: R Package gss. *J. of Statistical Software* 58(5): 1-25.
- Ishi, C. T., Sakakibara, K.-I., Ishiguro, H., and Hagita, N. (2008) A method for automatic detection of vocal fry. IEEE Trans. on Audio, Speech and Language Processing 16(1): 47-56.
- Kane, John, Drugman, Thomas, and Gobl, Christer (2013) Improved automatic detection of creak. *Computer Speech & Language*, 27(4):1028-1047.

References

- Keating, Patricia and Esposito, Christina (2007*) Linguistic Voice Quality*. UCLA Working Papers in Phonetics #105, 85-91
- Kim, S.-J. (1998) Positional effect on tonal alternations in Chichewa: phonological rules vs. phonetic timing. Proceedings of the Chicago Linguistics Society 34.
- Lee-Kim, S.-I., Kawahara, S., and Lee, S. J. (2014) The 'whistled fricative in Xitsonga: its articulation and acoustics. Phonetica 71(1) 50-81.
- Moisik, S. Lin, H., and Esling, J. (2013) Larynx Height and Constriction in Mandarin Tones, volume Eastward Flows the great river. City University of HK Press, 187-205.
- Myers, S. (1999) Tone association and F0 timing in Chichewa. Studies in African Linguistics 28: 215–239.
- Perkins, J., S. J. Lee. S. J., and J. Villegas. (to appear). An Interplay between F0 and phonation in Du'an Zhuang tone", in Proceedings of the 5th International Symposium on Tonal Aspects of Languages.
- Rose, Phil (1980) Thai-Phake Tones: Acoustic, Aerodynamic and Perceptual Data on a Tai Dialect wikth Contrastive Creak. Unpublished Manuscript.
- Wei, Qingwen, and Qin, Guosheng (1980) *Zhuang yu jian zhi* (A brief description of the Zhuang language). Beijing: Minzu chubanshe. [In Chinese]
- Xu, Y. (1998) Consistency of tone-syllable alignment across different syllable structures and speaking rates. Phonetica 55: 179–203.
- Xu, Y. (1999) F0 peak delay: when where and why it occurs. In J. Ohala (ed.), International Congress of Phonetic Sciences 1999. San Francisco: n.p., pp. 1881–1884.
- Yip, Moira (2002) *Tone*. Cambridge University Press.
- Zhang et al. (1999). Studies on Zhuang dialects. Sichuan Minzu Chubanshe. [In Chinese]

Acknowledgements

- We would like to thank Wuming Zhuang speakers. We also thank Prof. Wentao Gu for connecting us with the members of the linguistics community at Guangxi University: Prof. Tang, Heng Lin and Lina Jo.
- This research was partially funded by Kakenhi grant #15K16745 awarded to Jeremy Perkins.