

THE ROLES OF PHONATION AND F0 IN WUMING ZHUANG TONE

Jeremy Perkins

University of Aizu
Kamiiawase 90, Aizu-Wakamatsu, Japan
jperkins@u-aizu.ac.jp

Seunghun J. Lee

International Christian University
1-1-1, Higashicho, Koganei, Tokyo, Japan
seunghunlee@gmail.com

Julián Villegas

University of Aizu
Kamiiawase 90, Aizu-Wakamatsu, Japan
julian@u-aizu.ac.jp

ABSTRACT

This study reports phonetic measurements of the tonal system of Wuming Zhuang. While previous analyses have described Wuming Zhuang's tone contrasts using F0 only, this study also investigates reports of creaky phonation in Zhuang dialects. A composite acoustic algorithm is applied as a way to compute creaky phonation and is offered as an alternative method for linguists interested in measuring phonation from the acoustic signal. The results show that only one of four speakers exhibits evidence for contrastive creakiness, suggesting that F0 is the primary acoustic cue to contrastive tones in Wuming Zhuang, with creakiness playing a marginal role if any. These findings for Wuming Zhuang contrast with Du'an Zhuang, where creakiness plays a major role in the tonal contrast.

INTRODUCTION

Previous work on the tone system of Wuming Zhuang has divided the tone system into upper and lower registers based on F0 contours and on traditional Chinese description of tone (Wei & Qin 1980). Tones can be paired with similar F0 contours, differing in their register. For example, tone 1 [24] and tone 5 [35] both have rising F0 contours; tone 2 [31] and tone 4 [42] both have falling contours; and tone 3 [55] and tone 6 [33] both have level F0 contours. Tones 7 and 8 are complimentary distribution with tones 1 to 6; while tones 1 to 6 are only found in syllables with nasal codas or no coda (unchecked syllables), tones 7 and 8 are found in syllables with obstruent codas (checked syllables). Tones 7 and 8 also differ in their F0 contour according to the length of the vowel. This results in 10 distinct tones, however it is likely that the four checked tones are allotones of the six unchecked tones, a possibility we do not explore further.

The current study investigates whether pairs of tones with similar F0 contours may also involve phonation differences. In a tonal system with similar F0 contours, it is relatively more likely that other cues may aid in distinguishing between those tonal categories. As such, Wuming Zhuang is an interesting language to search for these differences. Furthermore, Perkins, Lee & Villegas (to appear) show evidence that in Du'an Zhuang, creakiness plays a major role in contrasting tones. This paper then allows for a comparison between the two dialects' tonal systems with respect to the role of phonation.

METHODS

Three female speakers (WZ01, WZ03 and WZ14) and one male speaker (WZ13) of Wuming Zhuang were recorded reciting a word list consisting of 197 words, with 5 repetitions each (985 tokens total). Words were selected from two sources (Wei & Qin 1980 and Zhang et al. 1999). Monosyllabic words were selected that met the following criteria, in order to avoid potential effects on F0 and spectral properties of the vowel:

- (1) Only monophthongal vowels were included.
- (2) Onset consonants were limited to alveolar, palatal and velar obstruents
- (3) Nasal codas were elicited but excluded from analysis as they induced creakiness.
- (4) Among checked syllables, only alveolar and velar stop codas were included.

Chinese orthography for each word was presented via Power Point slides. Uniquely randomized sequences of the slides were prepared for each speaker. In some cases, words were elicited that were for a selected monosyllable occurring in a longer multisyllabic word. These words were displayed in the form “X (as in XYZ)” where X, Y and Z were separate Chinese characters, but only X was to be spoken. Four practice items were selected and used for all participants with two of them being of the form “X (as in XYZ)” to allow participants to acclimate to this style of elicitation. Participants were given breaks prompted by the slide show after 200, 600 and 800 slides.

In addition to the presentation described above, participants also read word lists embedded within frame sentences. The frame sentence was “我正在读__这个词” (in English, “I am reading this word __ now”) with spaces separating the item from the frame sentences. The frame sentences confirmed that participants did not use a list intonation that interfered with F0 and/or phonation. The frame sentences were similarly displayed via slides in random order. They were read prior to the word list, with three repetitions per item. The analysis was not performed on the frame sentence recordings.

The recordings were made at the sound-attenuated booth at Guangxi University in Nanning, China. Participants were seated and wore a head-mounted Shure microphone connected via a Microdot-to-XLR cable to a Marantz Professional PMD661MKII handheld flash field recorder at a sampling frequency of 48 kHz. The slides were presented on a MacBook Air computer, and were advanced by the experimenter using the keyboard.

Following the recording session, one .wav file was created per token. In cases where a token was produced that differed from the intended pronunciation, it was excluded from analysis. Syllable rhymes were marked via TextGrids in Praat (Boersma & Weenink 2015) using increased intensity and appearance and disappearance of the vowel formants to mark boundaries. A script was then run that adjusted all the rhyme boundaries to the nearest zero-crossing. Octave errors, produced by the Praat script, 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.

The recordings were then analyzed in Praat for F0 and converted to semitones re. median F0/speaker. F0 measurements were taken every 10 ms by creating one “Pitch” object per token file and using the “Get Value in Frame” command in Praat. Creakiness was analyzed in Matlab, using the method described in Kane et al. (2013), which uses a composite of acoustic features to determine a probability of creakiness. This method was trained with actual judgments of creakiness and outperformed other similar recent models. Creakiness measurements were taken every 10 ms, outputting a probability of creakiness ranging from 0

to 1. The F0 and creakiness data were time-normalized and then fit to a smoothing cubic spline ANOVA model, as implemented in Gu (2014) using R.

RESULTS

Results of the analysis of F0 and creakiness in unchecked tones are shown in Figure 1 and Figure 2 respectively, with 95% Bayesian confidence intervals.

The F0 contours generally agree with Wei & Qin's (1980) account of Wuming Zhuang. Tones 2 and 4 are classified as [31] and [42], consistent with the findings in Figure 1. Tones 1 and 5 are classified as rising tones, [24] and [35] respectively by Wei & Qin. Our results confirm that F0 rises to a higher point in tone 5 than tone 1, but there is no difference in the initial F0 of tones 1 and 5. Tones 3 and 6, [55] and [33] respectively, have the greatest variation across speakers. Tone 3 has a slightly rising contour rather than a level contour as Wei & Qin report. Tone 6 can be accurately described as [33], but speaker WZ03 produces tone 6 with a lower F0 and a sharply falling-rising contour. The checked tones 7S, 7L, 8S and 8L are generally consistent with Wei & Qin's characterizations as [55], [35], [33] and [33].

There is larger variation between speakers in creakiness than in F0. Whereas the SS-ANOVA model fit well with F0 ($r^2 = 0.715$), the model for creakiness did not ($r^2 = 0.332$). Speaker WZ13 appears to differ from the other speakers in that he is generally creakier than the three female speakers, and has an earlier peak. Speakers WZ03 and WZ14 have nearly identical, negligible amounts of creakiness for all tones. Only speaker WZ01's tone 3 and WZ13's tone 5 show evidence of significantly increased creakiness.

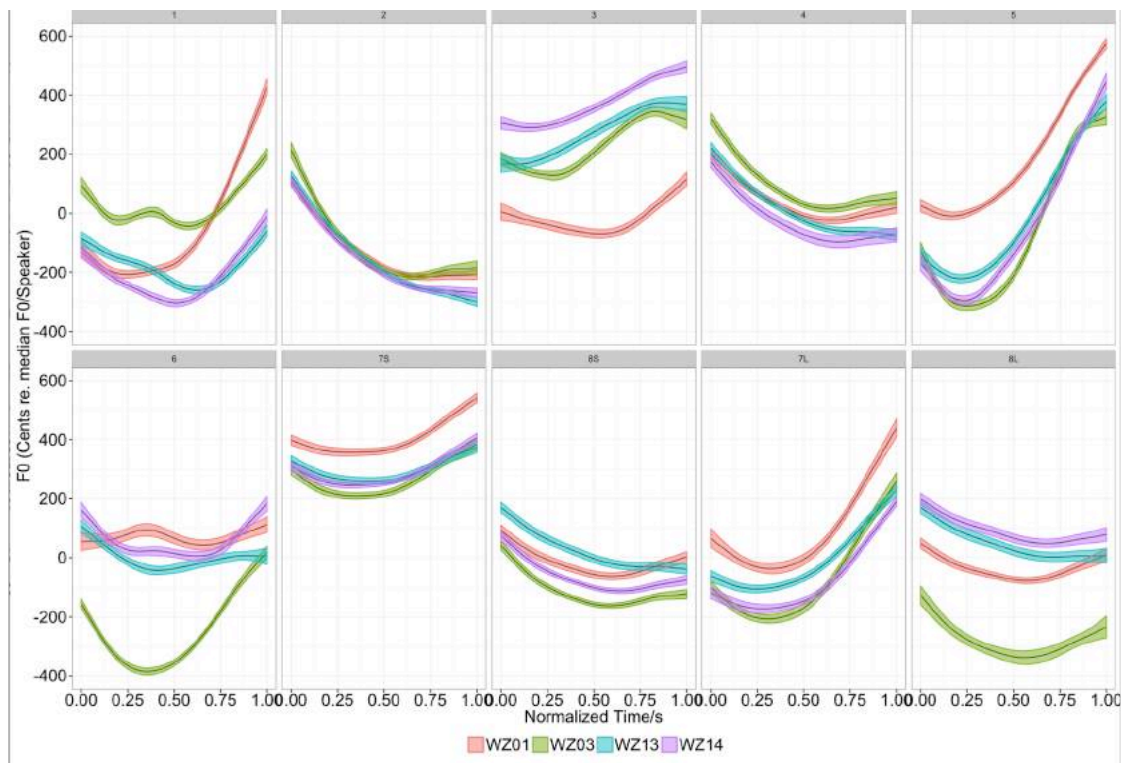


Figure 1: F0 contours for each speaker (different colours) by tone ($r^2 = 0.715$).

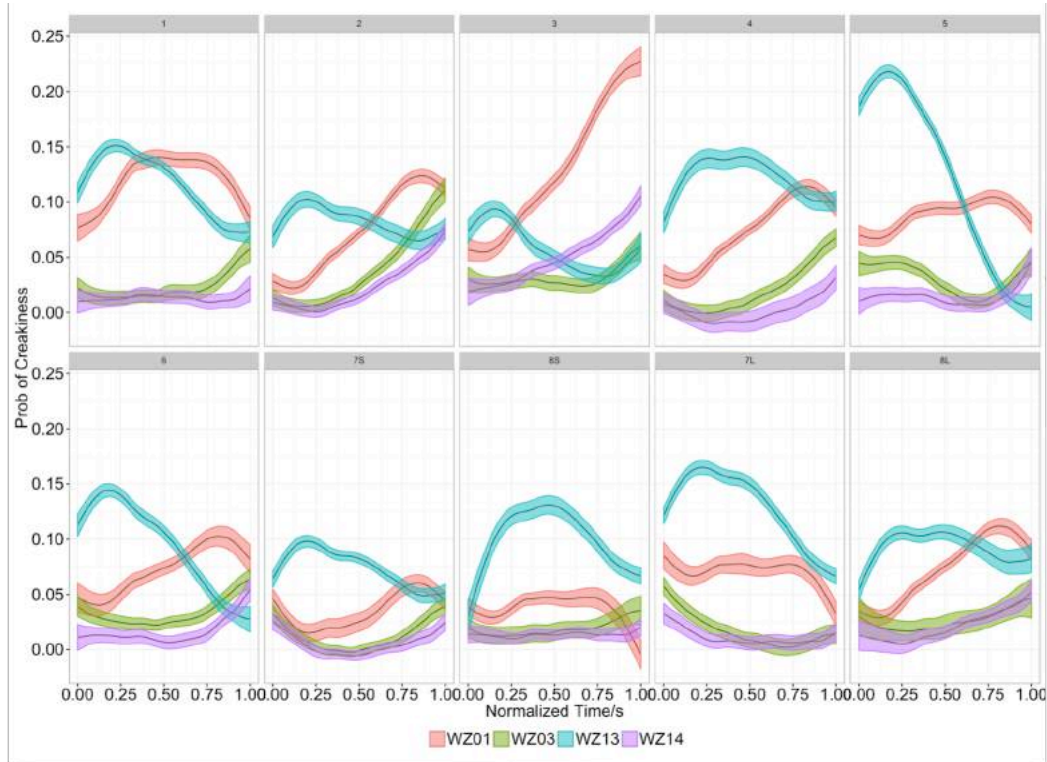


Figure 2: F0 contours for each speaker by Tone ($r^2 = 0.332$).

DISCUSSION & CONCLUSIONS

The results confirm that the tonal system of Wuming Zhuang is defined based on differences in F0 contours between tones, and not on creakiness as shown by the strong fit of the SS-ANOVA model for tone with F0, but not with creakiness. Furthermore, the variation between speakers was far greater with respect to creakiness than F0. We expect a larger degree of variation between speakers for a feature that is not contrastive in a language, as is the case here for creakiness. On the other hand, F0 did exhibit very little variation between speakers. This is true despite the fact that some tones have very similar F0 contours (tone 2 and 4 for example).

REFERENCES

- Boersma, P. and D. Weenink. (2016). Praat: doing phonetics by computer. Available [Mar. 2016] from www.praat.org.
- Gu, C. (2004). Smoothing Spline ANOVA Models: R Package gss. *J. of Statistical Software*, 58(5): 1–25.
- Kane, J., T., and C. Gobl. (2013). Improved automatic detection of creak. *Computer Speech & Language* 27(4): 1028–1047.
- 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*.
- Wei, Q. and G. Qin. (1980). 壮语简志 [Zhuang yu jianzhi] (A Brief Description of the Zhuang Language). Minzu chubanshe, Beijing. In Chinese.
- Zhang et al. (1999). Studies on Zhuang dialects. Sichuan Minzu Chubanshe. In Chinese.