様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成　３０　年　　６　月　２６　日現在

機関番号：　２１６０２
研究種目：　若手研究(B)
研究期間：　2015～2017
課題番号：　１５Ｋ１６７４５

研究課題名（和文）An Acoustic Typology of Creaky Voice

研究課題名（英文）An Acoustic Typology of Creaky Voice

研究代表者
Ｐｅｒｋｉｎｓ　Ｊｅｒｅｍｙ（Perkins, Jeremy）

会津大学・コンピュータ理工学部・准教授

研究者番号：３０７２５６３５

交付決定額（研究期間全体）：（直接経費）　　1,100,000円

研究成果の概要（和文）：この研究では、Wuming Zhuangでのきらきらした声の使用に焦点を当て、一対の音色にはきらきらとしたモーダルな音調のコントラストが含まれています。 また、新しいツール、心理音響的な粗さを導入しました。これは、ビルマのきらきらとした色調のクッキーをより確実に識別するために発見されました。 スペクトル測定に依存するCovarepアルゴリズムのような他の尺度は、ビルマのきらきら音を正確に予測しなかった。 最後に、Wuming Zhuangの話者に対して、知覚実験の三つ組を実行して、それらがどの程度かかりやすさ、F0および持続時間を使用して、対照音を区別するかを評価した.

研究成果の概要（英文）：This research focused on the use of creaky voice in Wuming Zhuang, finding that a pair of tones involve a creaky-modal phonation contrast. It also introduced a new tool, psychoacoustic roughness, which was found to more reliably identify creakiness in creaky and checked tones in Burmese. Other measures, such as the Covarep algorithm which relies on spectral measures, did not accurately predict creaky tones in Burmese. Finally, a triad of perception experiments were performed on Wuming Zhuang speakers to assess whether and to what extent they use creakiness, F0 and duration in distinguishing contrastive tones.

研究分野：　Phonetics

キーワード：　tone　phonation　Zhuang　Burmese　creakiness

１．研究開始当初の背景

This research investigated creaky voice, which is produced when the vocal cords are very stiff, so that the voice sounds rough. In some languages, creakiness is used the same way that different vowels are used in English and Japanese, to give words different meanings. For example, in the language Jalapa Mazatec, spoken in Mexico, the word *já,* means ‘tree,’ but if the vowel is creaky (*já)*, then it means ‘he carries’ (Ladefoged & Maddieson 1996: 317; Garellek & Keating 2011). To speakers of languages like English or Japanese, these words sound the same; however, Mazatec speakers recognize these words as sounding differently because they are sensitive to creakiness. Electroglottagraphy (EGG) is currently accepted as the best way to directly measure non-modal phonation; however many linguists do not have access to the equipment and training needed to collect EGG data. As a result, most linguists measure creaky voice acoustically, focusing on spectral properties of sounds. Creaky voice usually involves relatively more energy at higher frequency harmonics than modal voice. However there is no consensus currently on how to measure this “spectral tilt”. Linguists usually rely on a choosing one of a number of different measures. This research seeks to add a new acoustic measure of creakiness “psychoacoustic roughness” which relates directly to the perception system, and is particulary well-suited for cases where creakiness may be contrastive.

２．研究の目的

The purpose of this research was twofold. First, since most linguists do not have access to direct measures of glottal activity (e.g. electroglottography), this research uses new acoustic methods to confirm the presence or absence of creaky voice. Second, this research studied the tone system of two dialects of Zhuang, a Tai-Kadai language spoken in Guangxi province of China. Preliminary studies on a single speaker of Du’an Zhuang indicated that creakiness may be contrastive in that language, similar to Jalapa Mazatec and many other languages that employ creaky phonation contrastively.

One such other language is Burmese, which was used as a control case here, since it is known to have creaky tones. Field data collected from Zhuang and Burmese were analyzed using these two new acoustic methods, in order to clarify the tonal system of Zhuang and to show the value of these new methods of measuring creaky voice, for wider use among linguists. Finally, perception experiments were run on Zhuang speakers to confirm whether and to what extent F0 and creakiness play a role in tone recognition.

３．研究の方法

Research on creakiness has been mostly limited to speech production in understudied languages. This research project augments the already existing literature on production of creakiness by using two new algorithms to measure creakiness in the acoustic signal. The first algorithm utilized an objective measurement of psychoacoustic roughness as a proxy of creakiness, developed by Daniel & Weber (1997). The second algorithm was developed in Matlab by Kane et al. 2013, and is based on a composite of multiple acoustic cues. While Electroglottography remains the best direct method to measure creakiness, these two acoustic measures do not require any additional hardware or training, and are thus more accessible and easier to use for linguists without training on EGG methodology.

Psychoacoustic roughness is a sensory attribute related to rapid changes on the amplitude envelope of a sound (15 – 300 Hz). The rate and apparent extent of these changes is considered at different frequency bands. A predictor of creakiness that is based on roughness may allow linguists to have a better understanding of creaky phonation as a means of phonological contrast since it focuses on the reception of speech rather than the production side, unlike other acoustic measures (including Covarep).

The Covarep algorithm was used in addition; this algorithm determines the odds of a frame being creaky based on a cluster of acoustic features. The algorithm was based on observed correlations between these acoustic features and actual creakiness judgments on speech corpora containing creakiness.

Our research uses both of these methods to analyze creakiness in acoustic data from a language that is known to have creaky tones (Burmese, a control case). These results can then be compared to those from a language where there is no previous detailed phonetic work on the tone system (Zhuang), in order to determine whether some tones in Zhuang are in fact creaky.

Recordings were collected from field work on Zhuang done in Nanning, China at Guangxi University in December 2015 and on Burmese in Yangon in November 2016. Stimuli lists

were prepared for simple one-syllable words with an equal number of each tone, controlling for consonant sounds preceding and following the vowels and for the kinds of vowel sounds also. Six speakers of the Du'an dialect of Zhuang were recorded and eight speakers of the Wuming dialect of Zhuang were recorded. In Yangon, twelve Burmese speakers were recorded. The Zhuang stimuli consisted of 1,030 tokens (five repetitions of 206 words per speaker); the Burmese stimuli consisted of 78 words (1 repetition per word) per speaker. All speakers were in their 20s, with a roughly equal number of males and females for each language.

Finally, three perception experiments were run with speakers of Wuming Zhuang to confirm whether and to what extent speakers used F0 and/or creakiness and/or vowel duration to identify tones. Stimuli were chosen from the Wuming Zhuang recordings that were collected in December 2015. Variation in F0, creakiness (measured both by Covarep and by roughness) and duration was seen between speakers and this was capitalized on in the design of the experiment. Stimuli were chosen so that they had differing characteristics in F0, creakiness and duration, but the same tone. For each characteristic that differed in this way within a given tone, we could then assess whether that characteristic was relatively important in perceiving that tone by comparing the accuracy across speakers.

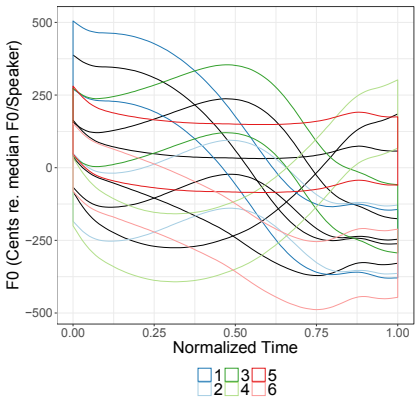
Three perception tests were performed on eight Wuming Zhuang speakers in March 2018. The first test was a forced choice test, where participants listened to a stimulus and were asked to choose which of four Zhuang words they heard. The words were written in Chinese and differed only on the tone, when possible, with one distractor added in cases where a four-way minimal quadruplet for tone didn't exist.

The second test involved the same set of stimuli as the first, but this time, the participants could freely type the word they heard using a Chinese pinyin keyboard.

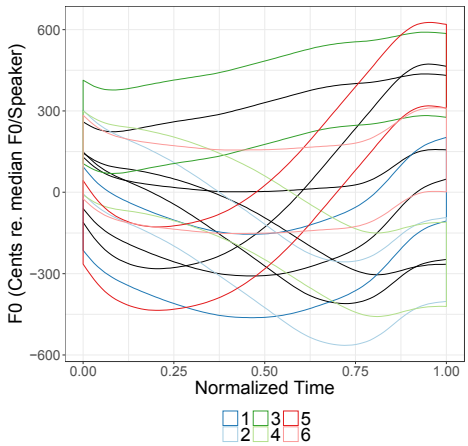
Finally, the third test played pairs of Zhuang words that either differed only on the tone, or were identical. Participants had to choose whether the two words were the same or different.

4．研究成果

Wuming Zhuang and Du'an Zhuang were found to have very different tone systems, which explains, in part, why they are mutually unintelligible. Analysis of F0 was done using SS-ANOVA with speakers as random intercepts. The following plot shows speaker-normalized F0 against normalized time for the six unchecked tones in Du'an Zhuang.



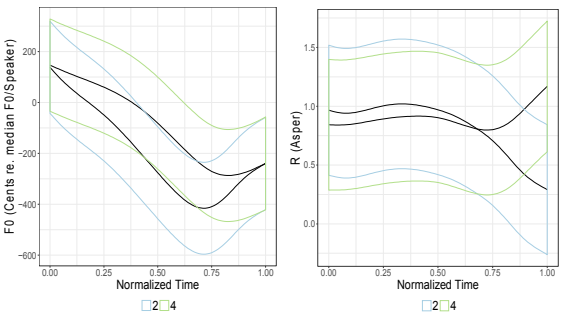
This can be compared to the following plot showing F0 in unchecked tones in Wuming Zhuang.

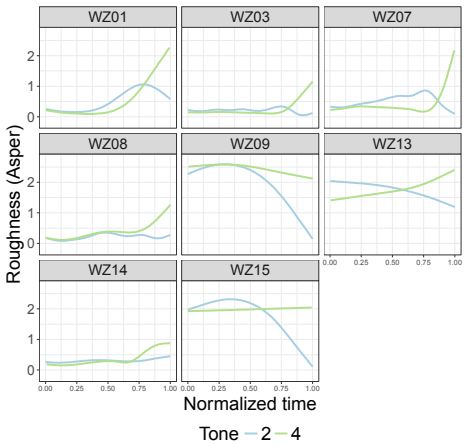


To give one example, tone 1 is a falling tone in Du'an Zhuang, but is a mid falling-rising tone in Wuming Zhuang. Comparing these plots, it's apparent that the tonal systems are completely different.

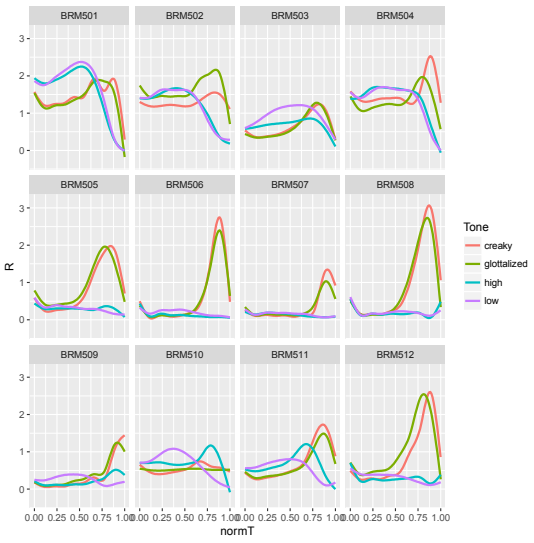
Regarding creakiness, there was evidence of a creaky-modal phonation contrast between tones 2 and 4 in Wuming Zhuang, but no evidence of contrastive creaky phonation in Du'an Zhuang. This result differs from the findings in Perkins, Lee & Villegas (2016) where a single Du'an Zhuang speaker in the United States had a modal tone 2 and a creaky tone 4.

While tone 2 and tone 4 had nearly identical falling F0 contours in Wuming Zhuang, the psychoacoustic roughness profiles diverged late in the vowel, as shown below (F0 on the left, Roughness on the right). Results shown for roughness separated by speaker are shown below. Units for roughness are Aspers; relatively higher values indicate creakiness.

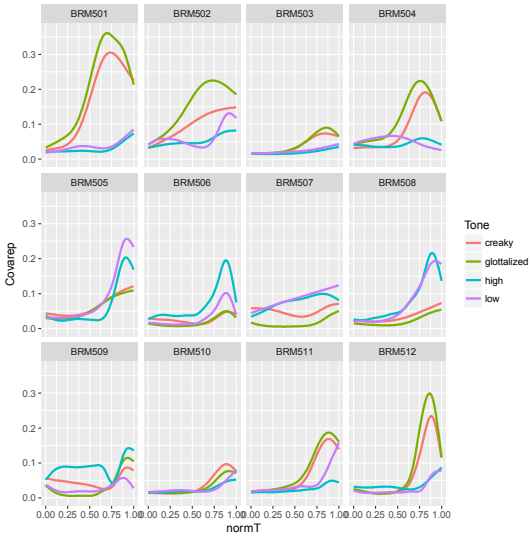




Analysis of tones in Burmese shows that psychoacoustic roughness more accurately predicts creaky and checked tones than any of the spectral tilt measures or creakiness probability from the Covarep algorithm. The results are shown by tone for each speaker separately below when reciting word lists. Notably, roughness is higher in creaky and checked (glottalized) tones late in the vowels for most speakers.



The results for creaky tones in Burmese are compatible with those from Wuming Zhuang in that both have late increases in roughness. These results imply that Wuming Zhuang tone 4 is creaky, similar to Burmese creaky and checked tones. They also demonstrate that psychoacoustic roughness is a valuable additional tool that linguists can use to identify creaky phonation acoustically. For comparison, we include the results of the Covarep algorithm, which uses spectral tilt to calculate an output probability of creakiness. The results, shown below, do not visibly correlate well with creaky and checked (glottalized) tone in most speakers.



These results are currently being prepared for publication. Results for the March 2018 perception experiments are not yet available, as they are currently being analyzed.

5．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔**雑誌論文**〕（計 6 件）

"Creakiness in Tone in Two Dialects of Zhuang" (in prep.) Jeremy Perkins, Julián Villegas, and Seunghun J. Lee.

"Automatic prediction of creaky voice with psychoacoustic roughness" (in prep.) Julián Villegas, Jeremy Perkins, and Seunghun J. Lee.

"Psychoacoustic roughness as a measure of creakiness in two dialects of Zhuang" (2017). Jeremy Perkins, Seunghun J. Lee and Julián Villegas. Proc. of the 2017 Seoul Int'l. Conference on Speech Sciences, pp. 111-112.

"Psychoacoustic roughness as a proxy of creakiness in White Hmong" (2017). Julián Villegas, Jeremy Perkins and Seunghun J.

Lee. Proc. of the 2017 Seoul Int'l. Conf. on Speech Sciences, pp. 74-75.

"The Roles of Phonation and F0 in Wuming Zhuang Tone" (to appear). Jeremy Perkins, Seunghun J. Lee and Julián Villegas. To appear in Proc. of the 22nd Himalayan Languages Symposium.

"An Interplay Between F0 and Phonation in Du'an Zhuang Tone" (2016). Jeremy Perkins, Seunghun J. Lee and Julián Villegas. Proc. of the 5th Int'l. Symposium on Tonal Aspects of Languages, pp. 56-59. http://isca-speech.org/archive/TAL_2016/pdfs/13-Perkins.pdf

〔学会発表〕（計 4　　件）

The 336th Regular Meeting of the Phonetic Society of Japan, Dec. 2, 2017 at Showa Women's University, Tokyo, Japan; Jeremy Perkins "Phonation in the Tone System of Zhuang"

2016 Annual Summer Conference of the Linguistic Society of Korea: Its 60th Anniversary meeting, Jun. 17, 2016; Seunghun J. Lee, Jeremy Perkins and Julián Villegas "The roles of phonation in F0 in Zhuang"

Invited Talk: International Conference on Phonetics and Phonology, Sep. 25-27, 2015 at Keio University, Mita campus, Tokyo, Japan; Seunghun J. Lee, Jeremy Perkins and Julián Villegas "Acoustic methods in the study of tone; applications to Du'an Zhuang and Burmese"

25th Annual meeting of the Southeast Asian Linguistics Society, May 27-29, 2015 at Payap University, Chiang Mai, Thailand; Jeremy Perkins, Seunghun J. Lee and Julián Villegas "OCP effects in suffixes with Burmese creaky tone"

〔図書〕（計　0　件）

〔産業財産権〕

○出願状況（計　0　件）

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況（計　0　件）

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等

６．研究組織
(1)研究代表者

パーキンズ　ジェレミー（Perkins Jeremy）

会津大学・コンピュータ理工学部・准教授

研究者番号：*30725635*

(2)研究分担者
（　　）

研究者番号：

(3)連携研究者
（　　）

研究者番号：

(4)研究協力者
（Lee, Seunghun J.）
International Christian University, Div. of Arts & Sciences, Associate Professor