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

令和　５ 年　 ６ 月 １６ 日現在

機関番号：２１６０２
研究種目：若手研究
研究期間：2019～2022
課題番号：１９Ｋ１３１６２

研究課題名（和文）Psychoacoustic roughness as a measure of glottalization in consonants

研究課題名（英文）Psychoacoustic roughness as a measure of glottalization in consonants

研究代表者
Ｐｅｒｋｉｎｓ　Ｊｅｒｅｍｙ（Perkins, Jeremy）

会津大学・コンピュータ理工学部・上級准教授

研究者番号：３０７２５６３５

交付決定額（研究期間全体）：（直接経費）　　1,000,000円

研究成果の概要（和文）：韓国語とタイ語からデータを収集し、タイ語と韓国語の両方で音響心理学的粗さが声門化子音を識別できることを示した。しかし、この結果は韓国語よりもタイ語の方がより強固であり、他の方法では韓国語の方がより大きな効果が得られました。最後に、機械学習法を用いて、一般的に使用されている音響測定値でモデルを学習させました。この方法は、言語学者が複数の音響測定値を用いた生産研究で使用するのに有望な方法であることが示されました。その結果、声調開始時間が緊張子音を他と区別する主要な尺度であること、声門形成は緊張子音をレニスや吸気子音と区別するのに必要ない副次的な特徴であることが示されました。
翻訳：DeepL.

研究成果の学術的意義や社会的意義
This research showed that acoustic measures that were traditionally used for voice quality in vowels can also identify glottal constriction in consonants (in particular, psychoacoustic roughness). Also, machine learning was used to assess which acoustic measures can distinguish groups of sounds.

研究成果の概要（英文）：Data was collected from Korean and Thai that showed psychoacoustic roughness can identify glottalized consonants in both Thai and Korean. However, this result was more robust in Thai than in Korean, where other methods yielded larger effects for Korean. Finally, machine learning methods were used, training a model on commonly used acoustic measurements. This method was shown to be a promising method for linguists to use in production studies with multiple acoustic measures. It confirmed previous studies on Korean that f0 is the primary acoustic measure that distinguishes aspirated and lenis consonants. It also showed that voice-onset-time is the primary measure distinguishing tense consonants from others, showing that glottalization is a secondary feature, and not necessary to distinguish tense consonants from lenis or aspirated ones. These results will be presented at a top-level international phonetics conference, with associated publications.

研究分野：Phonetics

キーワード：Psychoacoustic roughness　glottalized consonants　spectral tilt　Thai　Korean　machine learning　random forest

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

Whereas English and Japanese only have two kinds of consonant sounds that differ regarding the voice setting (voiced, like "b" and voiceless, like "p", for example), other languages may have different numbers of such categories. For example, Thai and Korean both have *three* kinds of consonants that differ in their voice setting. In Thai, there is a consonant sound that is in between voiced and voiceless (e.g., "p^h" - "p" - "b"). In previous research, this intermediate unaspirated "p" sound and the voiced "b" sound have been cited as also involving a constriction in the glottis (or larynx). Likewise, Korean also has three kinds of consonant sounds (tense, lenis and aspirated), each differing in the glottal setting. Korean tense consonants are similar to the unaspirated Thai "p", in that they are voiceless and also glottalized. One problem that is difficult for linguistic researchers studying the sound systems of these languages is that consonant sounds are generally made with the mouth closed, and so the resulting acoustic signal is rather empty of content. Electroglottagraphy is one way to solve this problem, as it uses an apparatus directly in contact with the larynx, directly measuring electrical impedance associated with muscular activity there. Identifying these consonants from the acoustic signal is still possible though.

There are aspects in which consonants and tones (pitch) of vowel sounds are linked. First, historically, consonant sounds that involve differences in glottalization sometimes change over time so that those differences move onto the vowel, affecting the pitch or tone of the vowel, and vice versa. Second, the phonetic effect of a glottalized consonant at the start of an adjacent vowel is nearly identical to that of a vowel spoken with creaky voice, and so the same methodologies used to identify creaky voice can then be used to identify glottalization in consonants. Fortunately, there is much past research on the use of acoustic analysis to study tone and phonation (voice setting - creaky or breathy voice, for example) in vowels. Consonant sounds overlap a little with adjacent vowel sounds and during that period of overlap, glottal gestures originating from the consonant can be measured acoustically using methods traditionally used to analyze vowels.

The most commonly used acoustic measure used to analyze vowel phonation is spectral tilt. Spectral tilt involves the difference between the intensity at a higher frequency in the sound spectrum subtracted from the intensity at a lower frequency. In regular voicing, there is usually a drop in the intensity of higher frequencies in the sound spectrum, whereas in creaky voice, there is relatively more energy at the higher frequencies. This difference has allowed research to use spectral tilt to distinguish regular voicing from creaky voice, for example.

Another acoustic method is psychoacoustic roughness, which has been used recently in identification of creaky voice in Zhuang, White Hmong, and Burmese, outperforming traditional spectral tilt measures in its identification of creaky tones. There is then a research gap in the analysis of glottalized consonants that can be filled by applying these methods traditionally used only in vowel analysis. Thai and Korean consonants here can provide a data set to fill this gap.

２．研究の目的

The purpose of this research was to investigate whether traditional methods used in phonation analysis of vowels could be applied to measure glottalization in consonants. In addition to spectral tilt, which is commonly used among linguists studying phonation, a new acoustic measurement method, psychoacoustic roughness, is also used and assessed as a substitute for spectral tilt in identifying glottalization in consonants. While not originally planned as part of this project, machine learning was applied here as a tool to explore which acoustic measures may be necessary or sufficient to distinguish sets of sounds from each other. As such, this research opened up some new methods available to phoneticians to analyze differences between sounds.

In addition to offering new methods for phoneticians, this research also addressed questions about whether unaspirated consonants in Thai are glottalized. Much past research has

claimed that there is secondary glottalization in these sounds, but it has not been confirmed. This research collected a large sample of recorded data from native Thai speakers, using established methods to assess whether these sounds are glottalized. As a control case for comparison, regular word-final stops and glottal stops were included in the recordings, since these sounds have been confirmed to be glottalized.

３．研究の方法

A production study was conducted, with recordings created of native speakers of Thai and Korean. Acoustic analysis was then performed, via both psychoacoustic roughness and traditional spectral tilt in vowels immediately following these glottalized consonants in Thai and Korean, assessing the ability of each to identify glottalization in consonants. While the general consensus is that the Korean tense consonants are glottalized, there is less consensus regarding the situation with the Thai consonants. The Korean results are thus taken as a way to validate the use of psychoacoustic roughness, and the Thai results can then potentially elucidate our understanding of whether unaspirated consonants really are glottalized in Thai.

In addition, a second study was performed on the Korean data using machine learning via the random forest technique. Eight acoustic measurements (f0, VOT, spectral tilt, psychoacoustic roughness, closure duration, frication duration, aspiration duration, and release duration) were used in training. This method identifies which of these acoustic measures form necessary and sufficient conditions for successful machine learning, offering an additional way to find potential cues that native speakers may use.

４．研究成果

The results showed that in both Thai and Korean, psychoacoustic roughness does successfully distinguish glottalized from non-glottalized consonants, but to different extents. In Thai, psychoacoustic roughness yielded larger effect sizes and was able to identify glottalization in stop consonants *following* vowels, which are known to have strong glottal constrictions. However, when looking at consonants that precede vowels, there was no evidence of glottalization in unaspirated consonants both based on spectral tilt and psychoacoustic roughness. Finally, it was noted that psychoacoustic roughness correlated inversely with tone (pitch) of the vowels. Thai is a tone language that uses pitch to distinguish meanings on vowels. As such, when a low tone is produced, rather than a high tone, this difference is used to create different words, much the same way vowels differ in Japanese or English. It was found that psychoacoustic roughness was higher during a low tone and lower during a high tone. This finding suggested that roughness should be used cautiously when measuring consonant glottalization in tone languages, since it correlates strongly with f0.

Meanwhile in Korean, psychoacoustic roughness measures confirmed that tense consonants are glottalized. Unlike Thai, the effect sizes for psychoacoustic roughness were smaller than spectral tilt in distinguishing tense from lenis and aspirated consonants. This research differed from previous work in that it also included affricates and fricatives, in addition to stops. Both psychoacoustic roughness and spectral tilt measures showed that tense fricatives are glottalized; however, tense affricates saw much larger effect sizes for spectral tilt, clearly showing glottalization, while the effect size for psychoacoustic roughness was relatively smaller.

Finally, a machine learning experiment using the random forest method found that spectral tilt and psychoacoustic roughness were not necessary in distinguishing Korean tense consonants. This suggests that glottalization is a secondary feature for tense consonants. Psychoacoustic roughness was relatively less important than spectral tilt for the tense series, but was surprisingly found to be more important than spectral tilt in distinguishing the lenis and aspirated consonants. The reason for this is likely because lenis consonants always precede a low tone in Korean. Just as in Thai, psychoacoustic roughness is higher in vowels with low tone, and so this may explain its ability to identify lenis stops. Finally, this experiment confirmed previous findings in Korean that f0 (pitch) was the main acoustic measure distinguishing the lenis and aspirated consonants, but that voice-onset time (VOT) was the main acoustic measure distinguishing tense stops from the others. Interestingly, in cases where VOT was unavailable, other similar measures acted as proxies: release (frication

+ aspiration) duration distinguished tense affricates and frication duration distinguished tense fricatives.

５．主な発表論文等

〔雑誌論文〕　計0件

〔学会発表〕　計3件（うち招待講演　0件／うち国際学会　2件）

１．発表者名 Jeremy Perkins
２．発表標題 Acoustic Measurement of Laryngeal Constriction in Thai Consonants
３．学会等名 35th General Meeting of the Phonetics Society of Japan
４．発表年 2021年

１．発表者名 Jeremy Perkins, Dahm Lee, Seunghun J. Lee
２．発表標題 A Production Study of Korean Consonants
３．学会等名 20th International Congress of Phonetic Sciences（国際学会）
４．発表年 2023年

１．発表者名 Jeremy Perkins, Yu Yan, Dahm Lee, Seunghun J. Lee
２．発表標題 Machine Learning to Model the Three-Way Laryngeal Contrast in Korean
３．学会等名 20th International Congress of Phonetic Sciences（国際学会）
４．発表年 2023年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関