

Creaky voice as a function of tonal categories and prosodic boundaries

Jianjing Kuang¹

¹Department of Linguistics, University of Pennsylvania

kuangj@ling.upenn.edu

Abstract

This study looks into the distribution of creaky voice in Mandarin in continuous speech. A creaky voice detector was used to automatically detect the appearance of creaky voice in a large-scale Mandarin corpus (Sinica COSPRO corpus). As the prosodic information has been annotated in the corpus, we were able to look at the distribution of creaky voice as a function of the interaction between tone and prosodic structures. As expected, among the five tonal categories (four lexical tones and one neutral tone), creaky voice is most likely to occur with Tone 3 and the neutral tone, followed by Tone 2 and Tone 4. Prosodic boundaries also play important roles, as the likelihood of creak increases when the prosodic boundaries are larger, regardless of the tonal categories. It is also confirmed that the pitch range for the occurrence of creaky voice is 110 Hz for male speakers and 170 Hz for female speakers, consistent with previous small-scale studies. Finally, male speakers have a higher overall rate of creaky voice than female speakers. Altogether, this study validates the hypotheses from previous studies, and provides a better understanding of voice-source variation in different prosodic conditions.

Index Terms: voice quality, creaky voice, Mandarin, tone, intonation

1. Introduction

Creaky voice is a non-modal phonation that is often realized with irregular periodicity or very low f_0 . Depending on the acoustic details, creaky voice can be further categorized into a number of subtypes [1-5], but the differences among the subtypes are usually not linguistically meaningful. Thus, in this paper, we will only use the general term of creaky voice.

Creaky voice is related to various functions of speech communication, both at the linguistic level (e.g., prosodic variation) and paralinguistic level (e.g., emotion, style). In this paper, we will focus on the relationship between creaky voice and prosodic structures. There are several physiological reasons for creaky voice to occur: low f_0 , high level of adductive tension in the vocal folds, or low level of subglottal pressure [6]. As a result, creaky voice is sensitive to prosodic conditions, and is often associated with low-pitched and weakened positions (e.g., [7-15]).

There are at least several prosodic structures in Mandarin related to creaky voice. First, Mandarin has four lexical tones: Tone 1 (high level), Tone 2 (rising), Tone 3 (low/dipping), and Tone 4 (falling). Among the four lexical tones, Tone 3 is reported to occur frequently with creaky voice ([16-21]). However, previous studies (e.g., [16-19]) have shown that creaky voice is not exclusively tied to Tone 3; instead, it also can occur with Tone 4 and Tone 2, the tones containing low pitch targets. Creaky voice is a by-product of laryngeal

adjustment for producing low pitch, and functions as an enhancement cue for perceiving low pitch targets [19, 20]. Second, Mandarin has unstressed syllables: although most words in Mandarin are specified with lexical tones, a relatively smaller number of words contain the so-called neutral tone. Metrically, syllables with lexical tones are stressed (at least to some extent), whereas syllables with the neutral tone are always unstressed. Phonetically, syllables with the neutral tone usually show signs of weakening and lenition (e.g., [22]), such as reduced vowels, shorter duration, lower pitch, and possibly creaky voice. Further, it has been shown cross-linguistically that unstressed syllables are often produced with creaky voice [17,18]. Third, Mandarin has tone sandhi: in Mandarin, when two Tone 3 occur in a sequence, the first Tone 3 is realized as a rising f_0 contour, comparable to the contour of Tone 2. However, a corpus study [23] has shown that, in connected speech, the phonetic realization of Tone 3 sandhi is not the same as Tone 2. Would this affect the creakiness of sandhi syllables? Moreover, Kuang and Wang [24] showed that the likelihood of sandhi application is affected by the strength of the prosodic boundary between two adjacent Tone 3 syllables. Finally, Mandarin has several layers of prosodic grouping (e.g., syllables, prosodic words, small prosodic phrases, large prosodic phrases, in general terms), which are mostly indicated by the strength of breaks [25,26]. Final positions of big prosodic phrases are also common environments for creaky voice [4,8,11].

The goal of this paper is to examine the presence of creak as a function of prosodic conditions. We would like to explore how tonal categories and prosodic structures interact and affect the presence of creaky voice. Previous studies [16-19] concerning this topic only examined monosyllabic words read in isolation. In this present study, we would like to investigate this question with connected speech data.

2. Methods

2.1. Corpus data and annotations

The COSPRO corpus, collected by Academia Sinica of Taiwan [26], was used for this study. In this corpus, both segmental information and the indexes of prosodic boundaries are annotated. The results reported in this paper are from COSPRO_01. This corpus contains speech data from six Taiwan Mandarin speakers (3M, 3F), and each speaker reads 599 short discourses (between 1 to 180 characters in length). The recordings took place in soundproof chambers and were digitized at 16 kHz with 16-bit resolution. This corpus was designed to include all possible syllables in Mandarin (about 1300), the most frequently used 2–4-syllable lexical words, all possible segmental combinations and concatenations, and all possible tonal combinations and concatenations. After data cleaning, a total of 347,196 tokens (112,760 vowel intervals) were analyzed in this study.

Since the speech data of COSPRO_01 is Taiwan Mandarin, it should be noted that the phonetic details of tone production are slightly different from Beijing Mandarin. For example, compared to Beijing Mandarin, Tone 3 in Taiwan Mandarin is usually produced as a low falling tone, instead of a low dipping tone. In addition, the phonetic realization of the neutral tone in Beijing Mandarin is highly dependent on the preceding syllable; contrastingly, the phonetic realization of the neutral tone of Taiwan Mandarin is fairly congruent and independent of the phonetic context, showing a trend of phonologization [27].

The break index system in this corpus follows Tseng [25]’s prosody model, which proposes six levels of prosodic grouping. From the top down, they are defined as:

- Prosodic Group (PG), the end of which is often realized with a big pause and final lengthening and weakening; breath group (break index = B5)
- Breath Group (BG), the end of which is often followed by inhaling, coupled with final lengthening and weakening (break index =B4)
- Prosodic Phrase (PPh), the end of which has a clearly perceived pause (break index = B3)
- Prosodic Word (PW), the end of which is often followed by a slight tone or voice change (break index=B2)
- Syllable (SYL). Syllable boundaries have no perceptible pause or lengthening. (break index =B1)

Break indexes B2 – B5 were used for analysis. Since the corpus was prosodically annotated to the word level, B1 is not explicitly labeled. Therefore, in our analysis, B2 refers to “below the prosodic word level”, including both syllable and prosodic word boundaries.

In addition, in order to take tone sandhi into consideration, Tone 3 tokens that are followed by another Tone 3 (T3s hereafter) were coded as a separate category from Tone 3 tokens that are followed by other tonal categories.

2.2. Analyses

Two analyses were performed to investigate the distribution of creaky voice in Mandarin. The presence of creak was automatically detected by an automatic creak detector, and the relative creakiness was estimated by a number of voice measures.

2.2.1. Automatic creaky voice detection

A creaky voice detector [28,29] was used to automatically identify the presence of creak in the speech corpus. The algorithm combines the features developed by Kane and Drugman [2, 28, 29], which is based on the characteristics of the excitation peaks in linear prediction residual signals, and the features developed by Ishi [30], which focuses on pulse-to-pulse regularities. Classification of creak was done by artificial neural networks, and the output of the classifier was the binary decision of creak (0 for non-creak, 1 for creak) for every 10ms. We then calculated the mean probability of creak for each segment by averaging out the all the decision values (0s and 1s) over time intervals within each segment.

2.2.2. Voice measures

In the second analysis, we examined the acoustic space of voice quality as a function of tonal categories and prosodic breaks. Voice measures were automatically extracted by VoiceSauce [31], including: fundamental frequency (F0), measured using the STRAIGHT algorithm [32]; periodicity and relative noise ratio measure — cepstral peak prominence (CPP)[33], harmonics-to-noise ratio [34], and subharmonics-to-harmonics ratio (SHR) [35]; the strength of the first harmonic relative to higher frequency components, namely $H1^*-H2^*$ (difference between the first two harmonics) and $H1^*-An^*$ (H1 relative to the amplitude of the formants). To allow for cross-vowel comparison in voice quality, these spectral tilt measures were corrected for vowel formants [38].

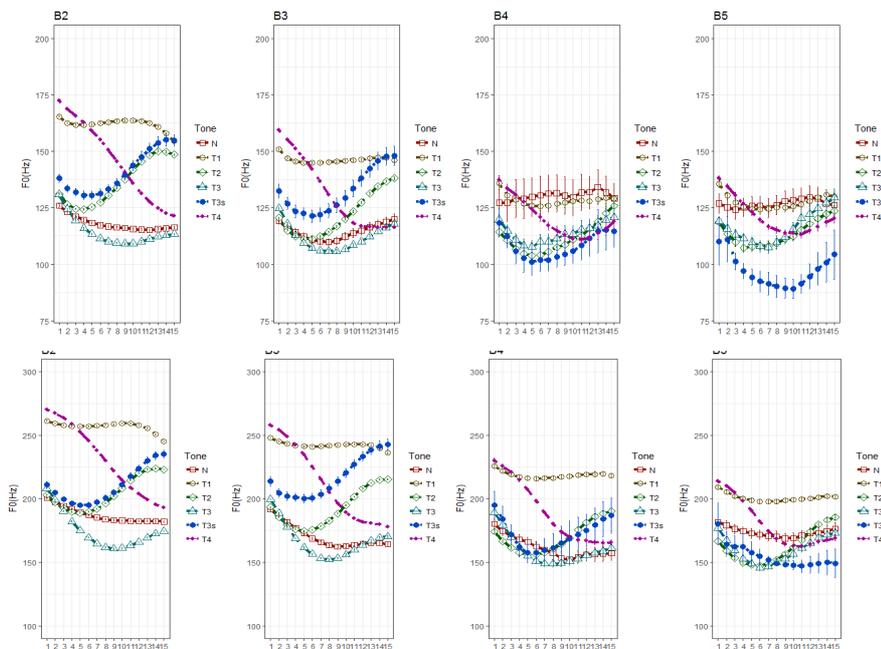


Figure 1: f_0 contours for tonal categories before different level of prosodic breaks. Top= male; bottom= female

3. Results and discussions

3.1. Probability of creak as a function of prosodic conditions

3.1.1. F0 contours of tonal categories

Because creaky voice in Mandarin is largely driven by pitch range [19], it is important to first examine the f0 contours of all six tonal categories (four lexical tones + neutral tone + T3s) before the different levels of breaks. As shown in Figure 1, the phonetic realization of tonal categories is affected by prosodic contexts. The overall f0 ranges of tonal categories are reduced as prosodic boundaries become larger, reflecting the global trend of f0 declination. The neutral tone is the tone least affected by prosodic breaks, as it is always located in the speaker's low-mid range regardless of the break indexes. This suggests that, although the phonetic realization of the neutral tone in Taiwan Mandarin is much less affected by the preceding tones, it still should be regarded as neutralized and unstressed, because it is always produced within the most comfortable and relaxing pitch range of the speakers.

Moreover, the overall shape of Tone 3 sandhi dramatically changes in different prosodic contexts. When the Tone 3 sequences cross the small prosodic boundaries such as B2 (within a word) and B3 (small phrase boundaries), Tone 3 sandhi has a clear rising f0 contour. Particularly, when the Tone 3 sequences cross the B2 boundary, the Tone 3 sandhi contour almost overlaps with Tone 2; but for B3, the phonetic difference between Tone 2 and Tone 3 sandhi is greater. When Tone 3 sequences cross a much larger boundary, such as B5 (utterance boundary), the f0 contour of Tone 3 sandhi is very different from Tone 2, and does not have a rising contour, which indicates that the sandhi rule rarely applies. Therefore, the likelihood of tone sandhi application heavily depends on the strength of the prosodic break between the two adjacent Tone 3 syllables. This result is consistent with [24], which observed that Tone 3 sandhi is more likely to take place for smaller breaks (e.g., word boundaries and syllable boundaries), and much less likely to occur for larger breaks; tone sandhi can sometimes cross intermediate phrase boundaries, but not big boundaries like intonational phrase or utterance boundaries.

3.1.2. Probability of creak

Figure 2 shows the probability of creak as a function of tones and breaks. Similar to [19], creaky voice is most likely to occur with Tone 3 and the neutral tone, followed by Tone 2 and Tone 4; Tone 1, as expected, is the least likely to creak. So, the presence of creak is indeed generally associated with low-pitch production as well as with vocal weakening. As for Tone 3 sandhi, its probability of creak depends on the prosodic breaks – the first Tone 3 in a two-syllable Tone 3 sequence is not likely to creak when the tonal sequence crosses smaller boundaries (e.g., B2 and B3), but it is much more likely to creak when the tonal sequence crosses larger boundaries (e.g., B4 and B5). The variation of creakiness for Tone 3 sandhi can be accounted for by the f0 variation shown in Figure 1. Recall that Tone 3 sandhi is more likely to apply across small prosodic boundaries, and less likely to apply across large boundaries. The probability of creak is clearly related to the realization of the f0 contours. As Tone 3 sandhi is sensitive to the strength of the prosodic boundary between two adjacent Tone 3 syllables, the likelihood of creak just varies

accordingly. The variation of Tone 3 sandhi thus again suggests that voice quality is co-varying with pitch ranges in Mandarin. In addition, Tone 3 sandhi overall exhibits different phonetic realization in both f0 and voice quality from Tone 2, supporting the notion that tone sandhi cannot simply be treated as Tone 3 turning into Tone 2.

Moreover, the strength of prosodic boundaries has a strong effect on the presence of creak. As the break index increases, the rate of creak also increases. Utterance-final (B5) syllables tend to creak regardless of tonal categories. We used *lme4* to perform a linear mixed effects analysis of the relationship between the probability of creak and tonal categories as well as break indexes. As fixed effects, we entered tone and break with interaction term into the model. As random effects, we had intercepts for speakers and segments for the effect of probability of creak. Since Tone 3 sandhi has a much smaller number of tokens, Tone 3 and Tone 3 sandhi are not separated in the statistical analysis. The *anova* function was used to estimate the overall significance of the factors (without breaking down into individual levels). The results show that both tone ($F[4,167] = 88.33, p < 0.01$) and break ($F[3,111807] = 665.73, p < 0.01$) have highly significant main effects on probability of creak, and the interaction between the two is also significant ($F[4,111779] = 17.49, p < 0.01$). Overall, the statistical results confirm our observation from Figure 2.

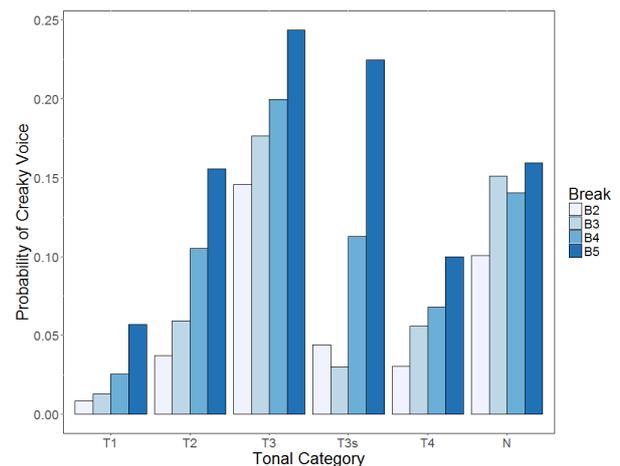


Figure 2: Probability of creak voice for tones and prosodic breaks

3.2. Acoustic space based on voice measures

Multidimensional-Scaling (MDS) was used to project the high dimensional space of tone and break categories into a two-dimensional space. All the measures mentioned in 2.2.2 are included in the calculation of distances among the categories. Figure 3 is the MDS acoustic space of voice quality of tonal categories in different prosodic contexts. The two dimensions together can account for about 70% of the variance. Correlation analysis indicates that dimension 1 is mostly correlated with spectral tilt measures. For example, $r = -0.85$ with H1*-A1*, and -0.8 with H1*-A2*; Dimension 2 is mostly correlated with noise measures, for example, 0.55 with SHR, and -0.35 with CPP. The direction of these measures suggests that the closer the upper left corner of the space, the creakier the voice. We have visualized this trend with an arrow in the plot. The overall distribution of the tone x break categories is

very consistent with the results from 3.1.2. As can be seen here, Dimension 1 represents the hierarchy of the prosodic boundaries, with higher levels of breaks on the left (creakier); and Dimension 2 is related to the tonal categories, with Tone 3 at the top (creakier) and Tone 1 at the bottom (less creaky). In addition, neutral tone categories are mostly along Dimension 1, suggesting that the creaky voice in the neutral tone is special in that it is mostly realized with spectral tilts, but very little with noise ratios. The results of this section overall validate the results from the creaky voice detector.

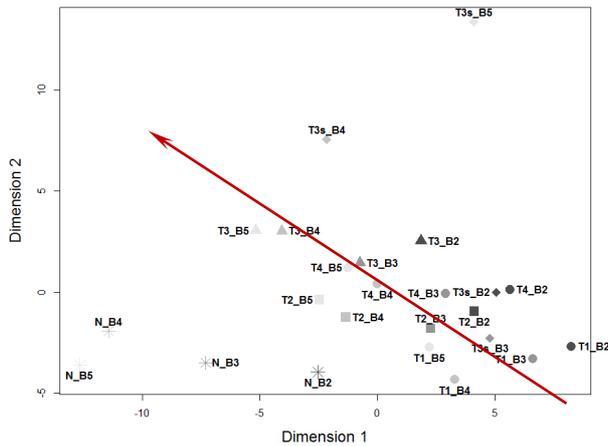


Figure 3: MDS acoustic space of tones and breaks

3.3. Gender difference

We are also interested to know whether there is any gender difference for the presence of creak. The preliminary results of gender difference can be seen in Figure 4. This figure suggests that male speakers generally have a higher rate of creak than female speakers. Since COSPRO_01 only has six speakers, and one male speaker has a much higher rate of creak than other speakers, the pattern can be driven by this outlier. Therefore, more data from a larger number of speakers will be needed to validate this observation. Nonetheless, the relative distribution of creak among the prosodic categories is fairly similar between the two genders, as confirmed by a spearman (rank) correlation test ($r = 1$). So different genders do not differ in how they use creaky voice linguistically.

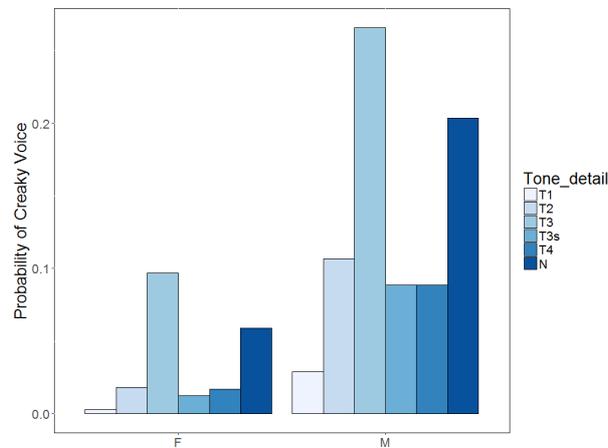


Figure 4: Gender difference in probability of creak

Finally, we would like to know in what the f_0 ranges creaky voice is likely to happen. We sorted out the creak regions (time intervals classified as 1) from the speech data of both genders, and extracted the f_0 values. Although the f_0 tracker would fail for a lot of tokens, it still returned an estimation of f_0 values during creaky voice. As shown in Figure 5, mean f_0 values of creaky voice are 170 Hz for female speakers and 115 Hz for male speakers, consistent with the numbers reported in [19].

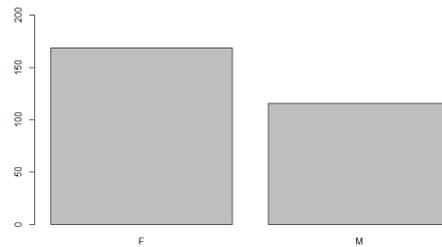


Figure 5: mean f_0 during creak for female and male speakers

4. Conclusions

This study examined the distribution of creaky voice as a function of tone and prosodic boundaries in Mandarin. Taking advantage of state-of-the-art technology, we were able to investigate this question with a large corpus of continuous speech. We found that both tone and the strength of prosodic boundaries exert a strong influence on the presence of creak. Specifically, among the five tonal categories (four lexical tones and one neutral tone), creaky voice is most likely to occur with Tone 3 and the neutral tone, followed by Tone 2 and Tone 4. The rate of creak is highly correlated with f_0 range. Prosodic boundaries also play important roles, as the rate of creak increases when the prosodic boundaries are larger, regardless of the tonal categories. We also found that Tone 3 sandhi differs from Tone 2 in terms of both f_0 contours and voice quality; and the rate of creak for Tone 3 sandhi depends on the realization of its pitch contour, which essentially depends on the level of break between the adjacent Tone 3 syllables. Finally, we provided some preliminary results for speaker variation. Although Male speakers overall have a higher rate of creaky voice than female speakers, the two genders do not differ in the way of using creaky voice. It is also confirmed that the pitch range for creaky voice to occur is about 110 Hz for male and 170 Hz for female, consistent with the numbers reported in our previous study. Altogether, this study validates the hypotheses from previous studies, and provides a better understanding of the voice-source variation in different prosodic conditions.

5. Acknowledgements

The study was supported by URF from University of Pennsylvania. I would like to thank John Bartlett for the assistance of data processing.

6. References

- [1] A. Batliner, S. Burger, B. Johne, and A. Kießling, "MUSLI:A classification scheme for laryngealizations," in *Proceedings of ESCA workshop on prosody*, Lund, 1993, pp. 176–179.
- [2] J. Kane, T. Drugman, and C. Gobl, "Improved automatic detection of creak," *Computer Speech and Language*, vol. 27, pp. 1028–1047, 2013.
- [3] P. Keating, M. Garellek, and J. Kreiman, "Acoustic properties of different kinds of creaky voice," in *Proceedings of the 18th Inter-national Congress of Phonetic Sciences, Glasgow*, 2015.
- [4] J. Slifka, "Some physiological correlates to regular and irregular phonation at the end of an utterance," *Journal of Voice*, vol. 20, pp. 171–186, 2006.
- [5] B. R. Gerratt and J. Kreiman, "Toward a taxonomy of nonmodal phonation," *Journal of Phonetics*, vol. 29, pp. 365–381, 2001.
- [6] J. Laver, *The Phonetic Description of Voice Quality* Cambridge: Cambridge University Press, 1980
- [7] J. Kreiman, "Perception of sentence and paragraph boundaries in natural conversation." *Journal of Phonetics*, vol. 10, pp. 163–175, 1982.
- [8] J. Pierrehumbert, "Prosodic effects on glottal allophones," in *Vocal fold physiology: voice quality control*, O. Fujimura and M. Hirano, Eds. San Diego: Singular Publishing Group, 1995, pp. 39–60.
- [9] L. Redi and S. Shattuck-Hufnagel, "Variation in the realization of glottalization in normal speakers," *Journal of Phonetics*, vol. 29, pp. 407–429, 2001.
- [10] D. H. Klatt & L. C. Klatt, "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *Journal of the Acoustical Society of America*, **87**, 820–857, 1990.
- [11] M. Garellek, "Perception of glottalization and phrase-final creak," *Journal of the Acoustical Society of America*, vol. 137, pp. 822–831, 2015.
- [12] K. J. Kohler, "Glottal stops and glottalization in German," *Phonetica*, **51**(1-3), 38-51, 1994
- [13] C. Mooshammer, "Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German." *Journal of the Acoustical Society of America*, vol. 127, 1047-1058, 2010
- [14] R. Ogden, "Turn transition, creak and glottal stop in Finnish talk-in- interaction," *Int. Phonetic Assoc.* vol. 31, pp. 139–152, 2001
- [15] L. Lancia, D. Voigt, & G. Krasovitskiy, "Characterization of laryngealization as irregular vocal fold vibration and interaction with prosodic prominence," *Journal of Phonetics* vol. 54, pp.80–97, 2016
- [16] A. Belotel-Grenié, & M. Grenié, "Phonation types analysis in Standard Chinese", *Proc. Spoken Language Processing* vol. 94, pp. 343–346, 1994
- [17] A. Belotel-Grenié, & M. Grenié, "Types de phonation et tons en chinois standard," *Cahiers de linguistique - Asie orientale*, vol. 26, pp. 249-279, 1997
- [18] A. Belotel-Grenié, & M. Grenié, "The creaky voice phonation and the organisation of chinese discourse," *In International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages* March 2004.
- [19] J. Kuang, *Phonation in tonal contrasts*, Dissertation, UCLA, 2013
- [20] R. X. Yang, "The Phonation factor in the categorical perception of Mandarin tones," in *Proceedings of ICPhS XVII*. Hong Kong, China, August 17-21, pp. 2204-2207, 2011
- [21] S. R. Moisiuk, H. Lin, & J. H. Esling. "A Study of Laryngeal Gestures in Mandarin Citation Tones Using Simultaneous Laryngoscopy and Laryngeal Ultrasound (slus)," *Journal of the International Phonetic Association*, vol. 44, pp. 21–58, 2014
- [22] B. Moore. "Phonetic observations on stress and tones in Mandarin Chinese," *The Journal of the Acoustical Society of America* **93**, 2297, 1993
- [23] Y. Chen and J. Yuan, "A corpus study of the 3rd tone sandhi in Standard Chinese," in *INTERSPEECH 2007 – 8th Annual Conference of the International Speech Communication Association*, August 27–31, Bonn, Germany, *Proceedings*, 2007, pp. 2749-2752
- [24] J. Kuang, & H. Wang, "T3 sandhi at the boundaries of different prosodic hierarchies," (in Chinese). *Journal of Chinese phonetics*, Vol. 1, pp. 125-131. 2006.
- [25] C. Tseng, S. Pin, Y. Lee, H. Wang, and Y. Chen, "Fluent speech prosody: framework and modeling," *Speech Communication, Special Issue on Quantitative Prosody Modelling for Natural Speech Description and Generation* vol. 46(3-4), pp. 284-309, 2005
- [26] C. Tseng, C. Chang, and Z. Su, "Investigation F0 Reset and Range in relation to Fluent Speech Prosody Hierarchy," *Technical Acoustics* vol. 24, pp. 279-284, 2005.
- [27] K. Huang, *A study of neutral-tone syllables in Taiwan Mandarin*, Doctoral dissertation, Honolulu, University of Hawaii at Manoa, 2012
- [28] T. Drugman, J. Kane, & C. Gobl, "Resonator-based creaky voice detection," in *INTERSPEECH 2012 – 13th Annual Conference of the International Speech Communication Association, September 9–13., Portland, Oregon, 2012*, pp. 1592-1595
- [29] T. Drugman, J. Kane, C. Gobl, "Data-driven Detection and Analysis of the Patterns of Creaky Voice," *Computer Speech and Language*, vol. 28, pp. 1233–1253, 2014.
- [30] C. Ishi, K. Sakakibara, H. Ishiguro, N. Hagita, "A method for automatic detection of vocal fry," *IEEE Transactions on Audio, Speech, and Language Processing* vol. 16 (1), 47–56, 2008
- [31] Y.-L. Shue, P. A. Keating, C. Vicens, and K. Yu, "VoiceSauce: A program for voice analysis," in *Proceedings of the International Congress of Phonetic Sciences, Hong Kong, 2011*, pp. 1846–1849.
- [32] H. Kawahara, A. de Cheveigné, and R. D. Patterson, "An instantaneous-frequency-based pitch extraction method for high-quality speech transformation: revised TEMPO in the STRAIGHT-suite," in *ICSLP-1998*, 1998, p. 0659.
- [33] J. Hillenbrand, R. A. Cleveland, and R. L. Erickson, "Acoustic correlates of breathy voice quality," *Journal of Speech and Hearing Research*, vol. 37, pp. 769–778, 1994.
- [34] G. de Krom, "A cepstrum-based technique for determining harmonics-to-noise ratio in speech signals," *Journal of Speech and Hearing Research*, vol. 36, pp. 254–266, 1993.
- [35] X. Sun, "Pitch determination and voice quality analysis using subharmonic-to-harmonic ratio," in *Proceedings of Acoustics, Speech, and Signal Processing (ICASSP)*, Orlando, FL, 2002, pp. I-333–I-336.
- [36] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, available at <http://www.R-project.org> 2014.
- [37] J. Kreiman and B. R. Gerratt, "Perceptual interaction of the harmonic source and noise in voice," *Journal of the Acoustical Society of America*, vol. 131, pp. 492–500, 2012.
- [38] M. Garellek, R. Samlan, B. R. Gerratt, and J. Kreiman, "Modeling the voice source in terms of spectral slopes," *Journal of the Acoustical Society of America*, vol. 139, pp. 1404–1410, 2016.
- [39] M. Garellek and P. Keating, "The acoustic consequences of phonation and tone interactions in Jalapa Mazatec," *Journal of the International Phonetic Association*, vol. 41, pp. 185–205, 2011
- [40] H. M. Hanson, K. N. Stevens, H.-K. J. Kuo, M. Y. Chen, and J. Slifka, "Towards models of phonation," *Journal of Phonetics*, vol. 29, pp. 451–480, 2001
- [41] M. Iseli, Y.-L. Shue, and A. Alwan, "Age, sex, and vowel dependencies of acoustic measures related to the voice source," *Journal of the Acoustical Society of America*, vol. 121, pp. 2283–2295, 2007.