3-30-2018

The Effect of Focus on Creaky Phonation in Mandarin Chinese Tones

Yaqian Huang
University of California, San Diego

Angeliki Athanasopoulou
University of Calgary

Irene Vogel
University of Delaware

The Effect of Focus on Creaky Phonation in Mandarin Chinese Tones

Abstract

Previous studies of the prosodic realization of focus in Mandarin Chinese show an expansion of the pitch range of lexical tones. It is less clear, however, whether focus affects the Creaky Phonation (CP) that often co-occurs with the Dipping third tone (T3), and to some extent, also with the Falling fourth tone (T4). This study investigates the effect of focus on the acoustic properties of the four Mandarin tones, and while it confirms the expansion of the pitch range under focus, it does not find that focus affects CP in T3; it only finds an effect of focus on CP in T4. Both the F0 and CP patterns are also considered in relation to the Functional Load Hypothesis, specifically, the relationship between the contrastive properties of a language and the manifestation of prominence.

The Effect of Focus on Creaky Phonation in Mandarin Chinese Tones

Yaqian Huang, Angeliki Athanasopoulou, and Irene Vogel

1 Introduction

While the four lexical tones in Mandarin Chinese are described as having different F0 contours, it has also been noted that the Dipping Tone (Tone 3) is often accompanied by Creaky Phonation. This may be due to the fact that the tone includes a portion with a low F0, which, in turn, causes creakiness to arise during the production of the tone. In fact, the Falling Tone (Tone 4) has also been found to exhibit some creakiness, although not as much as Tone 3, and this, too, may be due to the low F0 at the end of the tone contour. Alternatively, the presence of Creaky Phonation may serve as an enhancement of the Third, and possibly the Fourth, Tones, increasing their perceptibility and distinction from the each other, and the other tones. Since phonetic prominence typically enhances certain acoustic properties of focused elements, we address in this paper the question of whether placing Focus on words in Mandarin with Tones 3 and 4 increases their creakiness, thus further enhancing their tonal identity. Thus, following a brief introduction of the tonal properties of Mandarin (Section 2), we present the methodology of our experimental investigation of the effects of Focus on the different tones (Section 3). In Section 4, we present the results of our experiments, and in Sections 5 and 6, a general discussion and conclusions, respectively.

2 Mandarin Tones and Phonation

As is well known, Mandarin Chinese has four contrastive lexical tones: Tone 1 (High), Tone 2 (Rising), Tone 3 (Dipping), and Tone 4 (Falling). These tones are frequently represented schematically as in Figure 1.
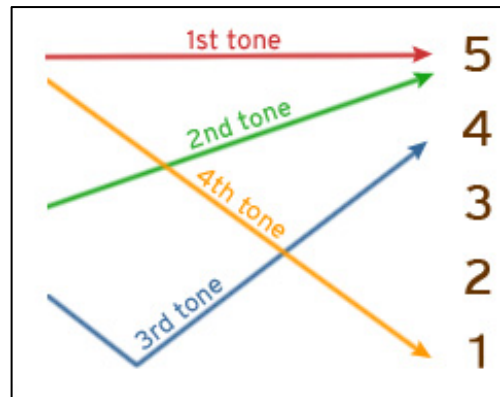

Figure 1: Mandarin Chinese Lexical Tones (ChinesePod 2017).

It has been documented (Belotel-Grenié and Grenié 1994, 2004, Davison 1991, Kuang 2013) that the Dipping Tone 3 (T3) is often accompanied by Creaky Phonation (CP), which may at least in part be due to the already low F0 in the middle of the tone. In fact, it has also been observed that the Falling Tone 4 (T4) may exhibit some CP (Belotel-Grenié and Grenié 1994, Kuang 2013), again likely related to the low F0, in this case, at the end of the tone. Moreover, with regard to T3, it has also been proposed that CP may, in fact, serve to enhance the perception of this tone (Yang 2011, 2015, Kuang 2013).

In stress languages, Focus frequently has the effect of further enhancing certain acoustic properties of the already enhanced main stressed vowel, and possibly additional material, of a word. In the absence of lexical stress in Mandarin, we consider here the possibility that Focus might increase the CP of T3, taking it as an enhancing property. We also consider the possibility that CP increases with T4 under Focus. In addition, since it has previously been observed that fo-

cus increases the pitch range in Mandarin, a possible F0 enhancement (Chen and Brauen 2006, Chen and Gussenhoven 2008, Jin 1996, Ouyang and Kaiser 2015, Xu 1999), we will also verify that this pattern is present in our corpus. In this case, we can determine that both CP and F0 properties may contribute simultaneously to the overall patterns of focus manifestation in Mandarin. Moreover, we examine all four tones to determine whether they behave similarly with regard to both CP and F0, although it is not anticipated that CP will be relevant for T1 and T2.

3 Experiment: Focus Effect on Phonation in Mandarin

As part of a larger cross-linguistic investigation of the acoustic properties of prominence, the Mandarin data have been collected in the same way as the data from other languages.

3.1 Hypotheses

To examine the effect of Focus on our stimuli, we compare the properties of our target vowels in two contexts, with and without focus. Specifically, we test Hypotheses 1-3 in (1), regarding the use of Creaky Phonation in the manifestation of Focus; in addition, we test a fourth hypothesis to confirm the previously noted increase in F0 range under Focus.

(1) Focus Manifestation Hypotheses
     a. Hypothesis 1: When focused, T3 will exhibit an increase in Creaky Phonation.
     b. Hypothesis 2: When focused, T4 will exhibit an increase in Creaky Phonation.
     c. Hypothesis 3: When focused, T1 and T2 will not exhibit an increase in Creaky Phonation.
     d. Hypothesis 4: Under focus, the tone space will exhibit an expanded F0 range.

In addition to using our data to test Hypotheses 1–4, also view our findings from a different perspective. That is, we consider the CP and F0 findings in light of a modified version of the Functional Load Hypothesis (FLH) that predicts a type of trade-off between the contrastive properties of a language and the manifestation of prominence (cf. Athanasopoulou and Vogel 2016, Vogel et al. 2016, Pincus et al. 2015). Specifically, it is predicted that a language will favor the use of acoustic properties that are not phonemic (e.g., duration in a language with contrastive vowel length) to express prominence (i.e., Focus, and Stress in languages where present), so as not to obscure lexical contrasts. While the use of F0 as a cue to Focus appears to conflict with the FLH, since F0 is a lexically contrastive property in Mandarin, an increase in CP would be in line with the FLH, since it is not typically viewed as a lexically contrastive property.

3.2 Experimental Design

3.2.1 Procedure

Speakers were tested individually in a lab of the School of International Studies at the University of International Business and Economics, Beijing. After training and practice, the speakers read dialogues and named objects provided on a computer in a PowerPoint presentation. The speech was recorded to the same computer with Praat, using a head-mounted microphone. The experiment lasted about 70-90 minutes.

3.2.2 Participants

The ten participants were monolingual Beijing Mandarin speakers from universities in Beijing (mean age = 21.83); however, here, only the recordings of the six male speakers are analyzed.
Although all the participants had studied other languages, none used them on a regular basis. None of the participants reported any history of reading, speaking or hearing impairment. They were paid $20 for their participation. Here are examples from English, as well as glossed examples.

3.2.3 Stimuli

The stimuli[1] were real three-syllable compounds, ideally of the form CVCVCV, familiar to the average Mandarin Chinese speaker, illustrated in Table 1. While some syllables had codas, the target vowels always appeared in a CV syllable. Each target vowel, /i, u, a/, appeared in 6 words with each of the 4 tones in each of the 3 syllables. All of the stimuli appeared in non-focus and focus contexts (see the next section), giving a total of 432 vowels per speaker, and a total corpus of 4320 vowels.

V1 = /a/		V1 = /i/		V1 = /u/	
word	**gloss**	**word**	**gloss**	**word**	**gloss**
b**ā** xiān zhuō	square table	d**ī** zī tài	modesty	f**ū** qī dàng	couple part-ner

Table 1: Sample Stimuli (Target = 1st syllable, high tone).

To avoid tonal confounds, the tones adjacent to the targets were strictly controlled so that they were as similar as possible to the targets. For example, a Rising T2, which begins relatively low, was preceded by a Falling T4, which ends relatively low; it was followed by T1 or T4 both of which begin high.

3.2.4 Carrier Dialogues

All stimuli were embedded in two carrier dialogues that place the target in a focused or a non-focused position, as seen in (1) and (2). The focused words are bolded; the target words are underscored.

(1) Focus Condition – Sample Dialogue:
   Q: *Lǎowáng biān shuō shénme biān xiězì?*
      'What did Laowang say while writing?'
   A: *Lǎowáng biān shuō **<u>shūfǎjiā</u>** biān xiězì.*
      'Laowang said "**calligrapher**" while writing.'

(2) Non-Focus Condition – Sample Dialogue:
   Q: *Lǎowáng shì biān shuō shūfǎjiā biān pàochá de ma?*
      'Did Laowang say calligrapher while making tea?'
   A: *Búshì, lǎowáng shì biān shuō <u>shūfǎjiā</u> biān **xiězì** de.*
      'No, Laowang said "<u>calligrapher</u>" while **writing**.'

The same requirements placed on the tones of adjacent syllables within the stimuli were applied to the adjacent syllables of target words and their carriers. Thus, for the words that have the target vowels at the beginning or at the end of the word, the carriers were slightly different so that the tone of the word before or after the target did not influence the tone of the target. For example, as seen in (1) and (2), where the target vowel in the first syllable of the stimulus, *shūfǎjiā*, has a high Tone 1, the preceding word, *shuō,* also has a high Tone 1.

In Mandarin, the interrogative word and corresponding answer are normally sentence-final. To avoid sentence final effects on the targets, however, the carriers included the collocation structure *bian…bian* 'while…' in order to place the stimuli in the middle of the response sentence. Note that in the Non-Focus Condition, the target word is placed before a focused word in the carrier sentence to avoid possible confounds from post-focal compression (Xu 1999). The dialogue slides, illustrated in Figure 2, alternated with filler slides with pictures the speakers had to name so

---

[1] The items conform to the requirements of a larger on-going project investigating the acoustic properties of prominence across languages (cf. Athanasopoulou and Vogel 2016, Vogel et al. 2016).

as to reduce the likelihood of the prosody of one dialogue influencing the next one.

Figure 2: Sample Slide.

3.3 Analyses

3.3.1 Acoustic Analyses

First, using Praat, all of the target vowels were manually segmented. Based on a combination of auditory assessment and visual confirmation, each vowel was coded as having Creaky Phonation if it sounded creaky to a native speaker (first author), and exhibited creaky properties in the spectrogram and waveform over at least half of the vowel's duration.

Subsequently, VoiceSauce was used for the acoustic measurements of the vowels, since it is more successful than Praat in measuring F0 with non-modal phonation. Since, according to Keating et al. (2015), the main acoustic correlates of CP are low F0, low Harmonic-to-Noise Ratio (HNR), low Cepstral Peak Prominence (CPP), and low H1-H2, we examined these properties, as well as the additional spectral tilt measure, H1-A1, in our data.

3.3.2 Statistical Analyses

One way 4*2 ANOVA followed by Bonferroni corrected pairwise t-tests, were used to evaluate the effects of tone, focus and their interaction on the distribution of creaky phonation and the phonation measures.

4 Results

The initial coding of the presence of Creaky Phonation in our corpus revealed that the male speakers exhibited much more CP than the females (i.e., 37% vs. 9% overall). We thus present here only the data from the male speakers since the question we are addressing specifically pertains to the use of CP.

4.1 Auditory/Visual Assessment of Creaky Phonation

Based on the auditory assessment of the recordings and the visual inspection of the spectrograms and waveforms, it was found that in both the basic Non-Focus Condition, and the Focus Condition, T3 exhibited by far the most CP, as shown in Figure 3.
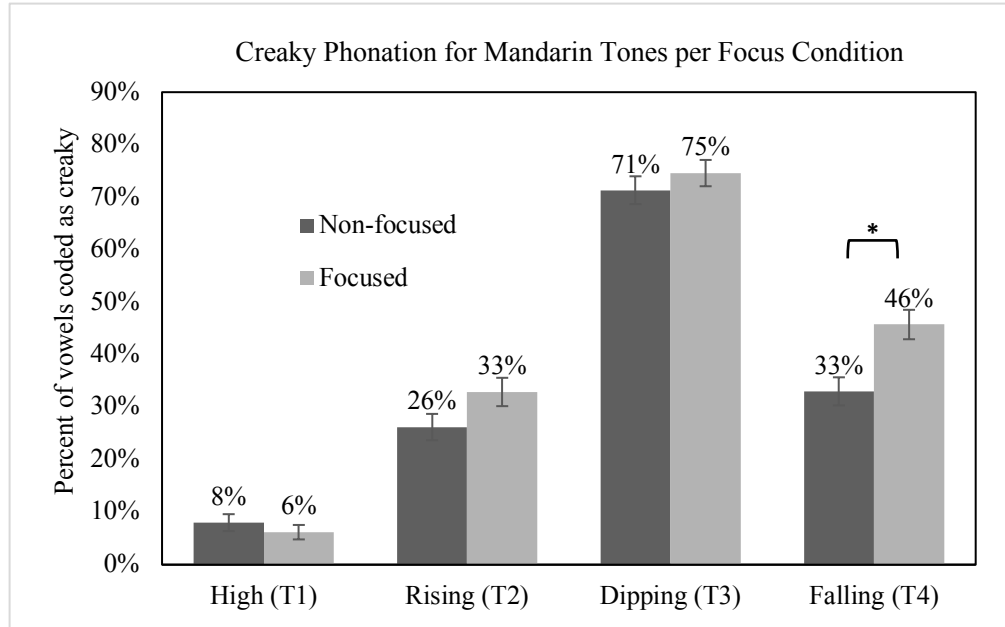
Figure 3: Distribution of Creaky Phonation per Tone; Focus and Non-Focus Conditions (error bars are SE based on the means of percentages across speakers).

A one-way ANOVA shows a main effect of focus [F (1, 9.674), p <.01], a main effect of tone [F (3, 251.694), p <.001], and an effect of interaction of tone*focus [F (3, 3.205), p =.022] on the use of CP across the four tones. As expected, T1 is essentially modal, and the minimal presence of CP is not increased under Focus. Also, as expected, T3 exhibits considerable CP; however, in this case, too, the rate of CP is not significantly increased under Focus. While a notable presence of CP was expected in T4, as can be seen in Figure 3, there was a similar amount of CP in T2 as well, a finding that was not expected based on previous literature. Bonferroni post-hoc t-tests of the differences between the mean percentages of CP in the Non-Focus and Focus Conditions only revealed a significance in T4 (p =.0011).

4.2 F0 Range

To examine the F0 range of the tones, and the possible effect of Focus on the range, each vowel was divided into three equal portions, and the mean F0 was calculated for each. Figure 4 shows the F0 contours of each tone in the Non-Focus and Focus Conditions.
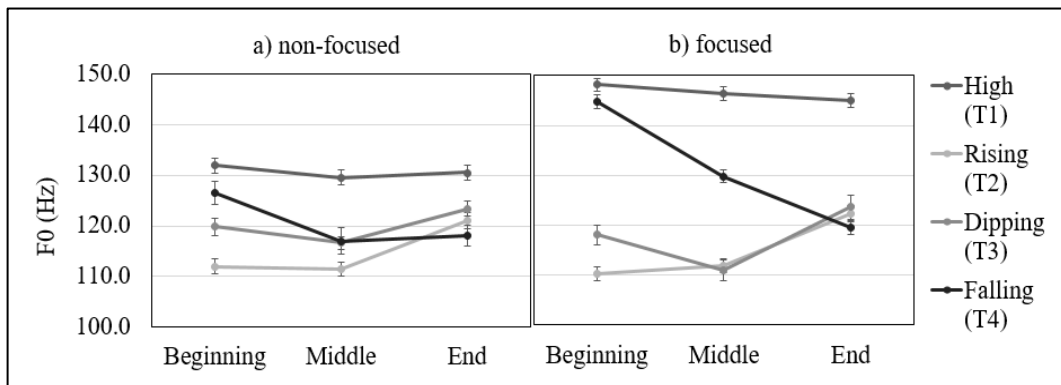

Figure 4: F0 Contours for Each Tone in Focus and Non-Focus Conditions (error bars are SE based on the means of F0s across speakers).

As can be seen in Figure 4, the overall F0 range is indeed expanded; however, this takes place mostly in one direction, towards higher F0 values. Moreover, the major changes only take place in the high portions of two tones: T1, which becomes higher throughout, and T4, which begins at a higher F0.

4.2.1 Additional Acoustic Properties

As noted, several acoustic properties have been found to correlate with CP, aside from lower F0, specifically, lower HNR, lower CPP, and lower H1-H2. We thus also examined these properties in our target vowels to determine to what extent they, in fact, correlated with the presence of CP observed in our data based on the initial auditory and visual assessments. Of the three properties, the only one that relatively correlated with the observed presence of CP was HNR. Figure 5 shows the mean Z-score of HNR values for each tone. The lower the HNR value, the creakier the vowel is. As can be seen, the HNR value is considerably lower for T3 than for the other tones; T1, with the most modal phonation, has the highest HNR value. As noted above, both T2 and T4 exhibit an intermediate presence of CP, and they correspondingly show intermediate HNR values.
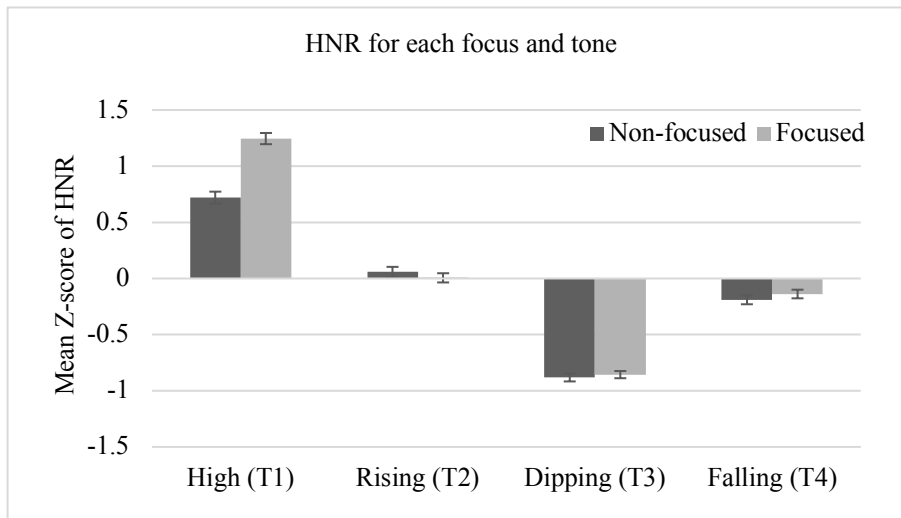

Figure 5: Harmonic-to-Noise Ratios for Each Tone in Focus and Non-Focus Conditions (error bars are SE based on the means of HNR Z-scores across speakers).

The result of one-way ANOVA shows a main effect of tone [$F_{(3, 634.08)}$, $p < .001$], focus [$F_{(1, 20.42)}$, $p < .001$] and the interaction of tone and focus [$F_{(3, 18.81)}$, $p < .001$] for the HNR values. Although the general pattern of creak parallels with the auditory and visual assessment, post-hoc Bonferroni t-tests doesn't show a focus effect in T4. Instead, the focus effect on HNR is only found in T1 ($p < .001$).

5 Discussion

The main goal of the present study was to investigate the possible influence of Focus on Creaky Phonation in Mandarin. While CP has previously been observed in general with Tone 3, and to some extent also Tone 4, the possible effect of Focus on CP has not previously been examined. Based on the auditory and visual perceptual coding performed by a native speaker of Mandarin, we observed, as in other studies (Belotel-Grenié and Grenié 1994, 2004, Davison 1991, Kuang 2013), considerable CP with T3, and some CP with T4. It was not expected, however, that we would also observe almost as much CP with T2 as with T4. As far as the effect of Focus on the presence of CP is concerned, for the most part, we did not find that the potentially enhancing CP property was further enhanced in T3 when Focused. The only tone to show a significant, though not a very large, increase in CP under Focus was T4. Thus, Hypothesis 1 was not confirmed, while

Hypothesis 2 was. Hypothesis 3 predicted that neither T1 nor T2 would exhibit an increase in CP under Focus; however, an increase was observed in T2, although it did not reach significance. Thus, Hypothesis 3 was confirmed. When the perceptual assessments of the presence of CP were compared with the acoustic properties that have been proposed as the main correlates of non-modal phonation, it was found that only one measure, HNR, reflected the distribution of CP as determined by the native speaker. That is, it showed a lower value for T3 than for the other tones, and it did not exhibit any consistent patterns of change under Focus. Though the pattern that T1 gets even more modal under focus could be of interest for future research.

The additional hypothesis regarding F0 range, Hypothesis 4, was partially confirmed in that the overall pitch range was considerably increased for T1 and T4, that is, where high F0s are involved. There was minimal expansion of the range at the lower end, that is, in the dipping portion of T3 and the low part of T4. It is possible that the lowest portion of the F0 range of the speakers was already being used for T3 and T4 from the outset, leaving little room for further lowering of the F0 value. This is consistent with earlier findings, where it was noted that the expansion of the pitch range could be observed especially towards the maximum F0 (e.g., Jin 1996, Shih 1988, Xu 1999).

Finally, with regard to the Functional Load Hypothesis, although CP would be available as an enhancing property to express Focus, since it is not contrastive in Mandarin, it was not exploited for this purpose. This neither confirms nor disconfirms the FLH. Instead, the fact that F0 was found to be used for expressing Focus prominence, as previously reported, at first glance appears to disconfirm the FLH, since F0, the main property of Mandarin tones, would not be expected to serve for the expression of prominence. Closer examination of the F0 patterns, however, reveals that a more nuanced interpretation of FLH is, in fact, confirmed. That is, since the main expansion of the pitch range is in the upper portion, where Focus causes the values to be even higher, the result is a further enhancement of the differences among the tones, rather than a loss of the contrasts due to the use of F0 as a prominence cue. In fact, the lower F0 values were not substantially raised under Focus, and thus did not jeopardize the lexical tonal contrasts.

6 Conclusions

In sum, while it was confirmed that T3 has the most concomitant presence of Creaky Phonation among the Mandarin tones, and that T4 has some presence of CP, as previously observed, it was also found that there was almost as much CP with T2 as with T4. It was predicted that if CP is an enhancing property of T3, it would show an increase under Focus, however, this was not observed. This suggests that CP in Mandarin is most likely primarily a reflection of the low F0 already associated with this tone, as opposed to constituting an "intentional" enhancement of the tone. An increase in CP under Focus for T4 was found; however, it is not clear that the extent of the difference is particularly noteworthy, although it was statistically significant. Of the various acoustic properties typically associated with non-modal phonation, only HNR showed any correspondence with the perceptually observed presence of CP. Finally, with regard to the F0 range, as in previous studies, it was found that the range did increase under Focus, but only at the higher end. This pattern conforms to a somewhat more nuanced interpretation of the Functional Load Hypothesis, since the use of F0 as a prominence cue did not compromise the tonal contrasts, even though F0 is otherwise used as a lexically contrastive property in Mandarin.

References

Athanasopoulou, Angeliki, and Irene Vogel. 2016. The acoustic manifestation of prominence in stressless languages. *Interspeech 2016* 82–86.

Belotel-Grenié, Agnès, and Grenié, Michel. 1994. Phonation types analysis in Standard Chinese. In *Proceedings of the 3rd International Conference on Spoken Language Processing* (ICSLP 1994), Yokohama, Japan.

Belotel-Grenié, Agnès, and Grenié, Michel. 2004. The creaky voice phonation and the organisation of Chinese discourse. In *International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages*. Beijing, China.

Boersma, Paul, and Weenink, David. 2017. Praat: doing phonetics by computer [Computer program]. Version 6.0.31, retrieved 21 August 2017 from http://www.praat.org/

Chen, Yiya, and Braun, Bettina. 2006. Prosodic realization in information structure categories in standard Chinese. In *Speech Prosody: 3rd International Conference*, ed. R. Hoffmann and H. Mixdorff, 54. Dresden: TUD Press.

Chen, Yiya, and Gussenhoven, Carlos. 2008. Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics* 36:724–746.

ChinesePod. 2017. ChinesePod.com. Retrieved from http://chinesepod.com/tools/pronunciation/section/17

Davison, Deborah S. 1991. An acoustic study of so-called creaky voice in Tianjin Mandarin. In *UCLA Working Papers in Phonetics* 78:50–57.

Gårding, Eva, Kratochvil, Paul, Svantesson, Jan-Olof, and Zhang, Jialu. 1986. Tone 4 and Tone 3 discrimination in modern standard Chinese. *Language and Speech* 29:281–293.

Jin, Shunde. 1996. An Acoustic Study of Sentence Stress in Mandarin Chinese. Doctoral dissertation. Ohio State University.

Keating, Patricia, Garellek, Marc, and Kreiman, Jody. 2015. Acoustic properties of different kinds of creaky voice. In *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow.

Kuang, Jianjing. 2013. Phonation in Tonal Contrasts. Doctoral dissertation. University of California, Los Angeles.

Ouyang, Iris Chuoying, and Kaiser, Elsi. 2015. Prosody and information structure in a tone language: an investigation of Mandarin Chinese. *Language, Cognition and Neuroscience* 30:57–72.

Pincus, Nadya, Athanasopoulou, Angeliki, Miller, Taylor, and Vogel, Irene. 2015. Evaluating acoustic measurements of creaky voice: A Vietnamese case study. *The Journal of the Acoustical Society of America* 138:1809.

Shih, Chilin. L. 1988. Tone and intonation in Mandarin. In *Working Papers of the Cornell Phonetics Laboratory* 3:83–109.

Vogel, Irene, Athanasopoulou, Angeliki, and Pincus, Nadya. 2016. Prominence, contrast and the Functional Load Hypothesis: An acoustic investigation. In *Dimensions of Phonological Stress,* ed. J. Heinz, R. Goedemans, and H. van der Hulst, 123-167. Cambridge: University Press.

Xu, Yi. 1999. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics,* 27:55–105.

Yang, Ruo-Xiao. 2011. The phonation factor in the categorical perception of mandarin tones. In *Proceedings of ICPhS XVII*, 2204–2207.

Yang, Ruo-Xiao. 2015. The role of phonation cues in Mandarin tonal perception. *Journal of Chinese Linguistics* 43:453–472.

Department of Linguistics
University of California, San Diego
La Jolla, CA 92037
huangyq@ucsd.edu
ivogel@udel.edu
angeliki@udel.edu