The source of creak in Mandarin utterances

Yuan Chai

Department of Linguistics, University of California San Diego, La Jolla, USA

yuc521@ucsd.edu

Abstract

In Mandarin, creaky voice quality frequently occurs with low tones such as Tone 3 and 4 (Belotel-Grenié & Grenié, 1994), due to the low F0 of these tones (Kuang, 2017a). Further, the final position in Mandarin utterances is frequently creaky (Zhang, 2016), perhaps due to declination in declarative sentences (Yuan & Liberman, 2014). In this study, we ask whether the presence of creak in final position of Mandarin utterances is also motivated by the pragmatic function of marking the finality. We conducted a production experiment to test the voice quality in initial, medial, and final positions of declarative sentences and three types of interrogative sentences: interrogatives ending with a "ma" particle, interrogatives ended with a "ba" particle, and bare interrogatives. The results show that, controlling for F0, the final position is more constricted than non-final positions. In declarative sentences, final position is *less* periodic than non-final positions, but nonetheless differs in voice quality for different types of interrogative sentences. Therefore, creak in final position of Mandarin utterances is not only a by-product of low F0, but is also motivated by the pragmatic function of marking the utterance finality and/or sentence type.

Keywords: Creak, F0, Tone, Prosody, Phonation

The source of creak in Mandarin utterances

1. Introduction

Creaky voice is typically characterized by low F0, in addition to irregular vocal fold vibration, and a more constricted glottis (Keating et al., 2015). As a result, in certain tonal languages, creaky voice has been found to co-occur with the low tones. Yu and Lam (2014) found that, among the six tones in Cantonese, creaky voice most frequently cooccurs with the lowest tone (21 or 11). In Mandarin, the dipping Tone 3 (214) is most frequently realized with creaky voice (Belotel-Grenié & Grenié, 1994). Low-rising Tone 2 (35) and high-falling Tone 4 (51) are occasionally accompanied by a creaky voice quality on their low F0 target, which occurs either at the beginning or the end of the contour (Belotel-Grenié & Grenié, 1994; Davison, 1991; Kuang, 2017a; Zheng, 2006). Kuang (2017a) studied the relationship between F0 and creaky voice in Mandarin tones. When the high-level Tone 1 was produced with a low F0 during whisper, they were produced with a creaky voice quality. When the low dipping Tone 3 (214) was produced with a high F0 in exclamation sentence, its creaky quality disappeared. These results indicate that creaky voice in Mandarin is a by-product of low F0.

However, creak can also be realized in the absence of a low F0. One alternative is to produce creaky vowels with a tense/pressed voice: the glottis is constricted while the F0 is high and regular (Keating et al., 2015). Tense voice is used when a language has high-pitched creaky tone: languages such as Jalapa Mazatec (Garellek & Keating, 2011), Mpi (Blankenship, 2002), Takhian Thong Chong (DiCanio, 2009), and Taiwan Min (Pan, 2017) all have high tones that are tense in quality, either because they are specified for that voice quality, or when a high tone cooccurs with phonemic creak..

The aforementioned languages demonstrate that, for tones realized with creak, the creaky voice can be either dependent of low F0 or independent of F0. The same is true for the creak associated with certain utterance positions. On the one hand, creak frequently occurs in utterance positions where F0 is low, such as the final position. Cross-linguistically, in declarative sentences, phrase-final position tends to have a lower F0 compared to non-final positions (*Brazilian Portuguese*: de Moraes, 1999; *German*: Fuchs et al., 2015; *American English:* Maeda, 1976; *Mandarin*: Shih, 1997; Yuan & Liberman, 2014). Meanwhile, phrase-final position tends to have a creakier voice quality cross-linguistically. Ní Chasaide and Gobl (2004) found that, in Irish and Swedish, creak occurs at the phrase-final position where the F0 is extra low. Similar cases of phrase-final creak have been found in English (Redi & Shattuck-Hufnagel, 2001), London Jamaican (Local et al., 1985), and Santa Ana Del Valle Zapotec (Esposito, 2003). The above studies did not test explicitly whether the phrase-final creak is dependent on low F0, but the co-occurrence of low F0 and creak in phrase/utterance-final positions suggests that the phrase-final creak is probably induced by the low F0.

On the other hand, the creak associated with certain utterance positions can be independent of F0. Epstein (2003) tested the voice quality of the initial and final positions in American English declaratives and yes-no interrogatives, finding that there was no significant relation between F0 and voice source parameters that affect voice quality. The initial and final positions of utterances were less modal than medial position for both declarative sentences (which had a low F0 target in final position) and yes-no interrogative sentences (which had a high F0 target in final position). This suggests that in American English, the non-modal phonation was used to denote phrase-boundary rather than a result of low F0 (Kreiman, 1982). Creak-like phenomena (e.g. glottalization) are also associated with phrase onsets. In English (Dilley et al., 1996; Garellek, 2013) and German (Pompino-Marschall & Żygis, 2010), word-initial vowels are more likely to be glottalized when in phrase-initial position than in phrase-medial position. Apart from phrase boundaries, creak has been found in other positions that are not necessarily associated with low F0: repairs, repetitions, and hesitations (Zhuang & Hasegawa-Johnson, 2008). Under the above circumstances, creak is used to denote pragmatic functions, rather than a by-product of low F0.

2. The Current Study

In Mandarin, the final position of utterances has been found to be creakier than non-final positions (Belotel-Grenié & Grenié, 2004; Kuang, 2017b; Zhang, 2016). But in Mandarin, F0 declines over the course the utterance (Shih, 1997; Yuan & Liberman, 2014), so that utterance-final position tends to have lower F0 than non-final positions. The purpose of the current study is to determine whether the creak in Mandarin utterances is driven by the low F0 in utterance-final position, whether it is affected by the pragmatic function of the utterance, or whether creak is affected by both position and pragmatic function. Different pragmatic functions are achieved by including both declarative and interrogative sentences, as well as different types of interrogatives.

In Mandarin, yes-no interrogative sentences (referred as interrogatives below) generally have a higher pitch register than declarative sentences (referred as declaratives below) (Shen, 1990; Yuan et al., 2002). Lin (2004) claimed that declaratives have a low boundary tone, whereas interrogatives have a high boundary tone. As a result, the F0 of declaratives declines as the utterance proceeds, while in yes-no interrogatives the F0 goes against declination. Thus, if creak is a by-product of the low F0 in sentence-final position, we expect the final position of declaratives to be creakier than non-final positions while the final position of interrogatives to be slightly creakier than the initial position, and less creaky than the medial positions. Alternatively, if creak is dependent on utterance-final position, we would expect that the final position of sentences is always creakier than non-final positions, regardless of the sentence type.

However, sentence type (e.g. a declarative vs. an interrogative) may also have an independent effect on the creak patterns across the utterances. Smith (2002) found that, in French, the voice quality at phrase-final positions differs between declaratives and interrogatives. The interrogative-final vowels were more periodic than declarative-final vowels, which was true for both interrogatives with a pitch fall (partial question, as “wh-”question in English) and those with a pitch rise (inverted question and echo question, as yes-no question in English). Smith suggested that sentence type had an effect on voice quality independent of F0. The current study intends to determine whether the difference in voice quality between French declaratives and interrogatives can extend to Mandarin. Chai (2019) tested the voice quality in different utterance positions of Mandarin declaratives and “ma” interrogatives and found that controlling for F0, in declaratives, the final position was less periodic than the non-final positions; while in “ma” interrogatives, the final position was more periodic. The current study replicates Chai’s study and tests two additional types of Mandarin interrogatives: “ba” interrogatives and bare interrogatives. In Mandarin, “ma” and “ba” interrogatives are constructed by adding a “ma” and “ba” particle to the end of declarative sentences respectively. Bare interrogatives have the exact structure of declaratives, but are marked by rising intonation. “Ma” and bare interrogatives have a rising pitch in utterance-final position, whereas “ba” interrogatives have a falling patch in the final position (Shen, 1990; Zeng et al., 2004). Controlling for lexical tone, in “ma” interrogatives, the F0 in final position (mean = 284.4 Hz) is slightly lower than initial position (mean = 299.8 Hz), and higher than the lowest F0 in the medial positions (mean = 258.9 Hz). In bare interrogatives, the F0 in final position (mean = 295.7 Hz) is close to initial position (mean = 298.8 Hz), and much higher than the lowest F0 in the medial positions (mean = 257.5 Hz) (Shen, 1990). The rising intonation in utterance-final position is more prominent in bare interrogatives than in “ma” interrogatives. Similar results have been found by Zeng, Martin, and Boulakia (2004). The F0 of “ba” interrogatives declines linearly, as with declaratives. The F0 in the final position of “ba” interrogatives is higher than the final position of declaratives, but lower than “ma” and bare interrogatives (Ruan, 2004).

Besides the differences in intonation pattern, “ma”, “ba”, and bare interrogatives further differ in their pragmatic meaning: the three types of interrogatives differ in terms of the degree of belief the speakers have in a given proposition. The speaker’s expectation in the answer to the question denoted by the interrogative sentence also differs among those three interrogative types (Li & Thompson, 1989; Liing, 2014). Semantically, interrogatives denote two proposition: a positive proposition (referred as *p* below) and a negative proposition (referred as ¬*p* below). Pragmatically, “ma” interrogatives are used when speakers does not have a bias towards either proposition (Example 1; Example 1 and 2 are retrieved from *The Lancaster Corpus of Mandarin Chinese* (McEnery & Xiao, 2004) and Example 3 from *Chinese Internet Corpus* (Sharoff, 2006)).

Example 1. “Ma” interrogative
俊芳问她：“你吃饭了吗？”她点头。
*Jun4fang1 wen4 ta1: “**ni3 chi1 guo4 fan4 le ma?**” ta1 dian3 tou2.*
Junfang (name) ask her you eat PERF meal PERF Q.MA she nod head
*“Junfang asked her: ‘**Have you eaten?**’ She nodded.”*

Pragmatically, “ba” interrogatives are used when speakers have a bias towards the positive proposition (i.e. a positive bias) (Example 2). They can be interpreted as equivalence to “don’t you think…?” or “wouldn’t you agree…?” (Li & Thompson, 1989).

Example 2. "Ba" interrogative

你很久没见过她了吧？我微笑，"十五年了。"
Ni3 hen3 jiu3 mei2 jian4 guo4 ta1 le ba? *Wo3 wei1xiao4, "shi2wu3 nian2 le."*
You very long not see PERF her PERF Q.BA? I smile, "fifteen years PERF"
*(someone asked,) "**You haven't seen her for a long time, right?**" I smiled, "For fifteen years".*

Bare interrogatives are used when speakers have a bias towards the negative proposition (i.e. a negative bias). It is used to express speakers' surprise or disbelief (Liing, 2014) (Example 3). The difference between bare interrogatives and "ma" & "ba" interrogatives is that the addressee of the bare interrogatives generally know the source of the speakers' insufficient belief, while there is no such constriction on the particle interrogatives (Liing, 2014). In other words, the propositions that bare interrogatives denote cannot be new information in the conversation. Otherwise, it will be pragmatically infelicitous to use a bare interrogative. Consequently, "ma" and "ba" interrogatives can be asked out of blue, while bare interrogatives require a context. Table 1 summarizes the main differences between different types of sentences in terms of their syntactic structure, intonation pattern, semantic meaning, and pragmatic function.

Example 3. Bare interrogative

你再夸奖我，我会不好意思脸红的。我继续说。
- *Ni3 zai3 kua1jiang3 wo3, wo3 hui4 bu1hao3yi4si1 lian3 hong2 de. Wo3 ji4xu4 shuo1.*
  you continue compliment me, I will shameful face red DE I continue say
  "If you keep complimenting me, I will feel shameful and blush."

怎么？一个大男人还会脸红？无香的语气很惊奇。
- *Zen3me?* ***Yi2ge4 da4nan2ren2 hai2 hui4 lian3 hong2?*** *Wu2xiang1 de yu3qi3 hen3 jing1qi2*
  what one man still will face red NAME DE intonation very surprised
  "'Really? ***A man like you blushes****?' Wuxiang (name) sounded very surprised."*

Table 1. Main differences across sentence type for declaratives/interrogatives
(Abbreviations: Decls: declarative sentences; Ints: interrogative sentences; I: Initial; M: Medial; F: Final)

	Decls	**"Ba" ints**	**"Ma" ints**	**Bare ints**
Syntactic structure	no particle	"ba" particle	"ma" particle	no particle
Intonation pattern (F0)	I>M>F	I>M>F	I>F>M	I=F>M
Semantic meaning	$\{p\}$	$\{p, \neg p\}$	$\{p, \neg p\}$	$\{p, \neg p\}$
Pragmatic function	100% belief in p	positive bias: 80% belief in p, 20% belief in $\neg p$	no bias: 50% belief in p and $\neg p$ each	negative bias: 20% belief in p, 80% belief in $\neg p$

2.1. Hypotheses

In sum, the research questions of this study are: 1) Controlling for utterance position and sentence type, what is the relation between F0 and the degree of creakiness? 2) Controlling for F0, is the final position of Mandarin utterances still creakier (more constricted and aperiodic) than non-final positions? 3) Does the effect of position differ between declaratives and interrogatives? 4) Does the effect of position differ among different types of interrogatives? Our hypothesis to the first question is that controlling for all other variables, low F0 is associated with high degree of creakiness. Our hypothesis to the second question is that controlling for F0, the final position will still be creakier than non-final positions. Our hypothesis to the third question is that the difference in voice quality pattern between Mandarin declaratives and interrogatives will resemble the difference found in French: the effect of utterance position on the degree of creakiness will be stronger in declaratives than in interrogatives, because we expect the final position in declaratives to be creakier than the final position in interrogatives.

For the fourth research question, we have two alternatives. If the difference in voice quality between declaratives and interrogatives is motivated by the difference in semantics, we expect the effect of position to be the same for all interrogatives despite their pragmatic and intonational differences. If the differences are motivated by the pragmatic function of each sentence, we expect the effect of utterance position to interact with each type of sentence. Since we hypothesize that declaratives will have a creakier end than interrogatives, we infer that the stronger the speakers' belief towards the positive proposition denoted by the sentence is, the creakier the final position will be. Under this hypothesis, the degree of creakiness in final position among those four types of sentences will rank as *declarative > "ba" interrogative > "ma" interrogative > bare interrogative*. For declaratives and "ba" interrogatives, utterance-final position will be creakier than non-final positions. The increase of creakiness from non-final positions to final position will be greater in declaratives than in "ba" interrogatives. For "ma" and bare interrogatives, utterance-final position will be more periodic than non-final positions. The increase of periodicity from non-final positions to final position will be greater in bare interrogatives than in "ma" interrogatives.

3. Experiment

3.1. Participants and Stimuli

We recruited 32 native speakers of Northern Mandarin (mean age = 19.9; mean age of starting learning English = 6.6; mean age of arriving US = 16.6; 25 women, 7 men). Northern Mandarin is the Mandarin dialect spoken in the north central-eastern region of China, including the provinces of Heilongjiang, Jilin, Liaoning, Hebei, Shanxi, Inner Mongolia, Beijing, Tianjin, and Shandong. We recruited speakers from the same dialect region to minimize the influence of dialect on intonation and voice quality. The speakers did not report any learning, speech, or reading disorder.

The stimuli consisted of sixteen experimental items. Within each experimental item, there are four conditions: declarative; “ba” interrogative, “ma” interrogative, and bare interrogative. Each trial included a dialogue with the final sentence in the dialogue being the target sentence. Within each target sentence, the target initial, medial, and final syllable are identical. All the target syllables all have high-falling Tone Four (51). The target segments include: /pau/, /pi/, /ta/, /ti/, /tu/, /twej/, /kwej/, /li/, /lje/, /lu/, /ɥi/, /ɥe/, /wej/, /wu/, /je/, /ku/. The criterion for selecting the target syllable was to avoid aspiration, nasal, and fricative onsets. Aspiration and fricatives may introduce extra noise to the vowel, which may affect the HNR measurement. Nasals usually have a resonance peak around 250 Hz (Stevens, 2000; Styler, 2017). The addition of nasal resonance will boost the intensity of H1 and H2 around 250 Hz and affect the accuracy of H1*–H2* values accordingly.

The target sentences were embedded in a dialogue because bare interrogatives cannot be asked out of the blue and require a context. In order to control for the length of stimuli among conditions, the other types of sentences were embedded in a context as well. In all target sentences, narrow focus was avoided in order to minimize the influence of focus on voice quality. The naturalness of all the dialogues was confirmed by four native speakers of Mandarin who have spoken the language for more than twenty years. Example 4 illustrates a sample experimental item. The complete stimuli can be found in the appendix.

Example 4.1. Declarative

Target	***wu4** li3*	*lao3 shi1*	*zuo4*	***wu4**li3*	*shi2yan4*	*shi2*	*chu1xian4*	*le*	*shi1**wu4***	
	Physics	*teacher*	*do*	*physics*	*experiment*	*when*	*appear*	*PERF*	*mistake*	

"The physics teacher made a mistake when doing a physics experiment."

Context 小李放学回家，告诉妈妈，物理老师被学校开除了。妈妈问他为什么，他说："物理老师做物理实验时出现了失误。"

"Xiaoli came back home from school and told his mom that their physics teacher was fired by school. His mom asked him why. He said, 'The ***physics teacher made a mistake when doing a physics experiment.****'"*

Example 4.2. Ba interrogative

Target	***wu4** li3*	*lao3 shi1*	*zuo4*	***wu4**li3*	*shi2yan4*	*shi2*	*chu1xian4*	*le*	*shi1**wu4***	*ba?*
	Physics	*teacher*	*do*	*physics*	*experiment*	*when*	*appear*	*PERF*	*mistake*	*Q.BA*

"The physics teacher made a mistake when doing the physics experiment, right?"

Context 小李在实验课上看到物理老师在做实验的时候发生失误，但是小赵坚持老师没有失误。小李便向另一个同学小高求证："物理老师做物理实验时出现了失误吧？"

"Li saw that during the lab session, the physics teacher made a mistake when doing an experiment. But Zhao insisted that the teacher did not make a mistake. Thus, Li asked another classmate, Gao, to confirm, ***'The physics teacher made a mistake when doing the physics experiment, did he?****'"*

Example 4.3. Ma interrogative

Target	***wu4** li3*	*lao3 shi1*	*zuo4*	***wu4**li3*	*shi2yan4*	*shi2*	*chu1xian4*	*le*	*shi1**wu4***	*ma?*
	Physics	*teacher*	*do*	*physics*	*experiment*	*when*	*appear*	*PERF*	*mistake*	*Q.MA*

"Did the physics teacher make a mistake when doing a physics experiment?"

Context 小李今天缺席了实验课。但是他听说这节课物理老师要做一个很容易失误的实验，便在课后问同学："物理老师做物理实验时出现了失误吗？"

*"Li was absent for today's lab session, but he heard that their physics teacher was going to do an experiment, and is quite likely to make a mistake during that experiment. Thus, Li asked one of his classmate who attended the lab, '***Did the physics teacher make a mistake when doing the physics experiment?***'"*

Example 4.4. Bare interrogative

Target	***wu4** li3*	*lao3 shi1*	*zuo4*	***wu4**li3*	*shi2yan4*	*shi2*	*chu1xian4*	*le*	*shi1**wu4**?*	
	Physics	*teacher*	*do*	*physics*	*experiment*	*when*	*appear*	*PERF*	*mistake*	

"The physics teacher made a mistake when doing a physics experiment??"

Context 小李印象中的物理老师业务熟练，做实验很有经验，从不失误。但是小赵告诉他今天物理老师在一堂实验课上出现了操作失误。小李非常惊讶，难以置信地重复道："物理老师做物理实验时出现了失误？"

"Li always thinks that his physics teacher is very skilled. His physics teacher has a lot of experience in experiments and never makes a mistake. But Zhao told Li that their physics teacher made a mistake when doing an experiment during a lab session. Li was very surprised. He could not believe it and repeated Zhao's words, 'The ***physics teacher made a mistake when doing a physics experiment??****'"*

Sixteen experimental items were split into four lists using a Latin-square design. Each condition (sentence type) was tested by four items in each list. Eight fillers were added to each list. Each speaker produced 48 dialogues in total [(16 target + 8 fillers) * 2 repetitions]. The stimuli were grouped into two blocks. The order of the dialogues was randomized within each block for each subject. The stimuli were displayed on a computer screen using Psychopy (Peirce, 2007). The recording took place in a sound booth. The productions were recorded at a 44.1 kHz sampling rate and a 32-bit quantization rate using a Shure SM10A head-mounted microphone and the Audacity recording software (2018).

3.2. Segmentation Criteria

The vowel of each target word was segmented from its onset to offset. When the target word begins with a stop, the vowel onset was marked from the release burst of the stop. When the target word begins with a glide, the vowel onset was marked when the amplitude of the first two formants increased significantly. When the target word was followed by a word with an obstruent onset, the vowel offset was segmented right before the stop closure or the onset of frication of the following syllable. When the target word was followed by a word with a sonorant onset, the vowel offset was segmented when the amplitude of the first two formants decreased significantly. When the target word was followed by a word without an onset (i.e. in a hiatus environment), the target vowel offset was labelled at end of formant transitions for the next vowel. When the target word occurred at the end of the utterance, the offset of the vowel was marked where the voice bar or the formant ended, whichever came first.

3.3. Measuring Creak

The presence of creak was assessed using a creaky voice detector (Drugman, Kane, Gobl, 2014). The creaky voice detector uses various acoustic properties including low F0, H1*–H2*, residual peak prominence, and pulse irregularities to determine whether creak is present. It makes a binary decision as to whether a sound is creaky every 10 ms. Kuang (2017b) showed that the detector is accurate for Mandarin roughly 80% of the time. Following Kuang (2017b), the probability of creak equaled the number of "creaky" decisions divided by the number of 10-ms intervals in the sound. For instance, if a given segment was 100 ms long and was marked as creaky in 10 positions, its probability of creak was 0.1.

The results of the creak detector were cross-validated by the acoustic properties of creak. F0, H1*–H2*, and HNR (harmonics-to-noise-ratio) of the vowel of each target word were calculated automatically using VoiceSauce (Shue, Keating, Vicenik, & Yu, 2011), which outputs a value every millisecond over a target interval. H1*–H2*, the difference in amplitude between the first and second harmonics (corrected for formant frequencies and bandwidths), is related to the degree of vocal fold constriction: a lower value of H1*–H2* is associated with a higher degree of constriction (Garellek, 2019). HNR measures the degree of noise: a lower HNR value represents a higher degree of noise. Both low H1*–H2* and low HNR are associated with creaky voice (Garellek, 2019): the low H1*–H2* during creak is due to increased constriction, and the low HNR is due to voicing irregularity.

4. Results

4.1. F0 trend within utterances.

To describe the F0 trend over the course of utterance, we compared the F0 in different utterances positions separately for each type of sentence. Tokens with an F0 larger than three standard deviations from the speaker's mean were excluded from the analyses. F0 was regressed on the utterance position with utterance position as the fixed effect. All the slopes and intercepts of the models were random at the subject and the sentence level. In declaratives, F0 generally declined linearly from the utterance-initial to the final position. The F0 in final position was lower than initial (β = -0.653, χ^2 = 47.315, p < .0001) and medial position (β = -0.273, χ^2 = 23.736, p < .0001). In "ba" interrogatives, the F0 in final position was also lower than initial (β = -0.742, χ^2 = 45.799, p < .0001) and medial position (β = -0.251, χ^2 = 15.566, p < .0001). In "ma" interrogatives, utterance-initial position generally had a higher F0 than the non-initial positions (β = -0.43, χ^2 = 35.755, p < .0001), but F0 did not differ between utterance-medial and final positions (β = 0.024, χ^2 = 0.272, p = .602). In bare interrogatives, utterance-initial position generally had a higher F0 than the non-initial positions (β = -0.399, χ^2 = 25.523, p < .0001), but F0 did not differ between utterance-medial and final positions (β = 0.081, χ^2 = 2.471, p = .116). Figure 1 shows the F0 tracks of four sentence types from the same experimental item produced by four female speakers. Figure 2 shows the average F0 by position in utterance for each sentence type.

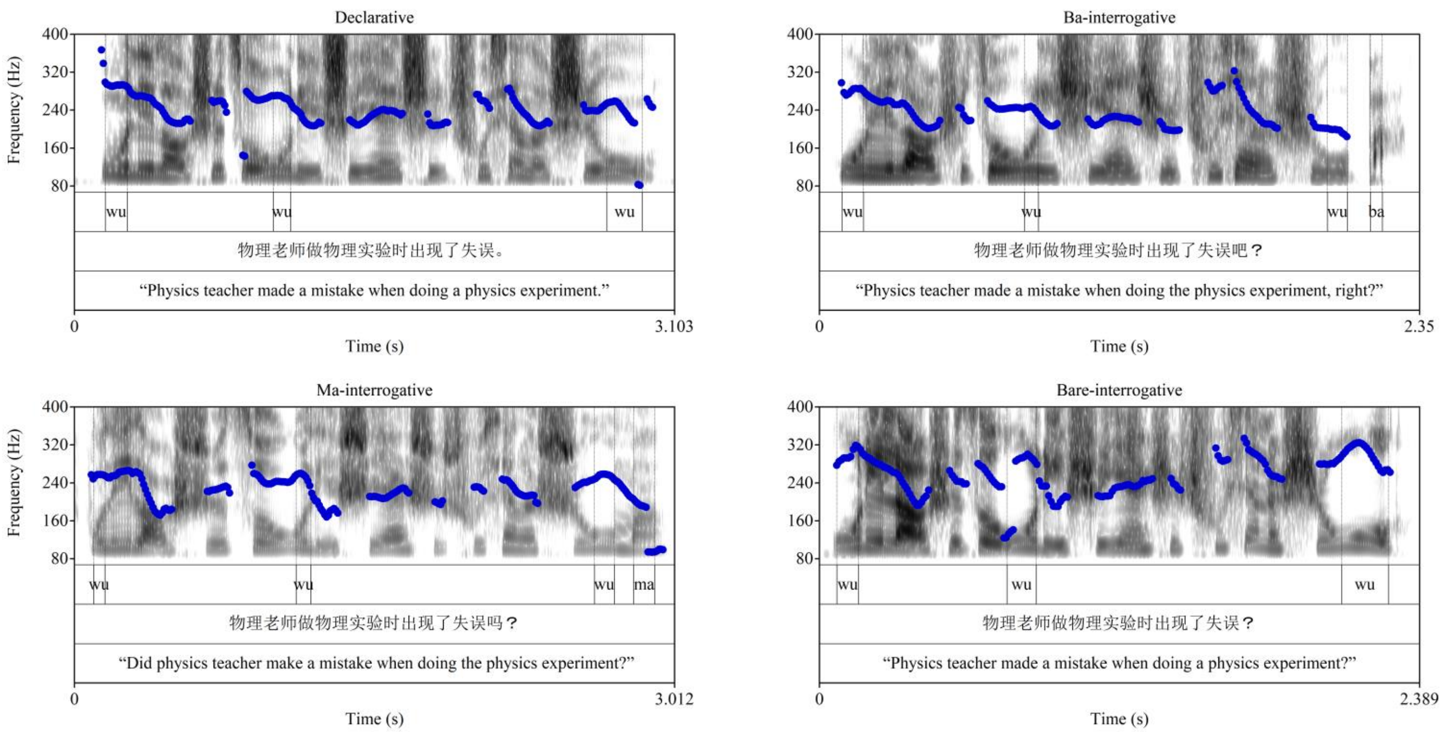

Figure 1. F0 track of four sentence types: declarative (top-left), ba-interrogative (top-right), ma-interrogative (bottom-left), and bare-interrogative (bottom-right).

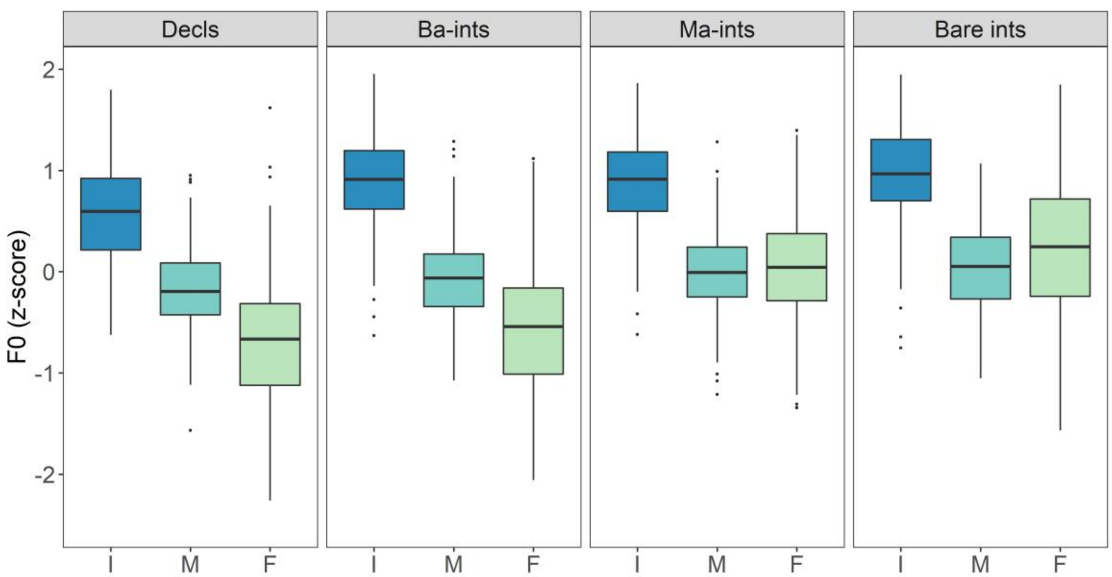

Figure 2. F0 in different utterance positions
(Decls: declaratives; Ints: Interrogatives; I: Initial; M: Medial; F: Final)

4.2. Probability of Creak

The probability of creak was calculated based on the results from the creaky voice detector. The probability of creak in each position in each sentence type is plotted in Figure 3. The probability of creak during the particle in "ma" interrogatives and "ba" interrogatives has also been included. Since declaratives and bare interrogatives both have three positions measured while "ma" and "ba" interrogatives both have four positions measured, they were split into two groups when testing the effect of position on probability. The declaratives were compared with bare interrogatives while "ma" interrogatives were compared with "ba" interrogatives. Within each group, the R code for regression is (1):

*(1) lmer(probability ~ position*type + (position || subject) + (position || sentence))*

Averaging across declaratives and bare interrogatives, the probability of creak was higher in final position than non-final positions ($\beta = 0.092$, $\chi^2 = 24.338$, $df = 1$, $p < .0001$). The difference in the probability of creak between final position and non-final positions was greater in declaratives than in bare interrogatives ($\beta = 0.144$, $\chi^2 = 22.605$, $df = 1$, $p < .0001$).

Averaging across "ba" interrogatives and "ma" interrogatives, the probability of creak during the particle was higher than the average of the other positions ($\beta = 0.318$, $\chi^2 = 54.009$, $df = 1$, $p < .0001$). The difference in the probability of creak between particle and other positions was greater in "ba" interrogatives than in "ma" interrogatives ($\beta = 0.332$, $\chi^2 = 70.895$, $df = 1$, $p < .0001$). For both "ba" and "ma" interrogatives, there was no significant difference between the "final" position (the position immediately before the particle) and the initial and medial position ($\beta = 0.002$, $\chi^2 = 22.605$, $df = 1$, $p = .774$).

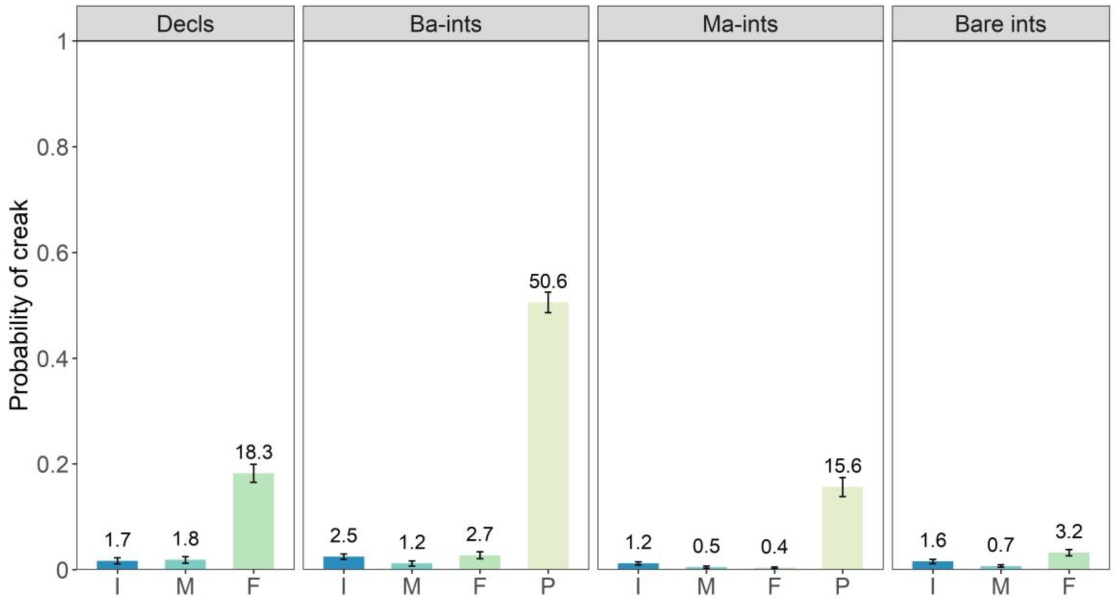

Figure 3. Probability of creak in different positions

4.3. H1–H2**

H1*–H2* was regressed on F0, utterance position and sentence type with F0, utterance position, and sentence type as the fixed effect. The slopes of F0 and utterance position, as well as the intercept, were random at the subject and sentence levels. The slope of sentence type was not included in the random-effects structure because the model did not converge. The R code for the model is (2):

(2) lmer(H1–H2* ~ (F0 + position) * type + (F0 + position || subject) + (F0 + position || sentence))*

All the interactions in Model (2) were not significant ($\chi^2(9) = 6.676$, $p = .671$). This indicates that the effect of F0 and position on H1*–H2* is consistent across sentence type. Since the interactions did not improve the model fitness, they were dropped from the model. The R code for the reduced model is (3):

(3) lmer(H1–H2* ~ F0 + position + type + (F0 + position || subject) + (F0 + position || sentence))*

The results of Model (3) showed that F0 was positively correlated with H1*–H2* ($\beta = 0.2$, $\chi^2 = 9.634$, $df = 1$, $p = .002$). The lower the F0 was, the lower the H1*–H2* was. For all four types of sentences, controlling for F0, the final position had lower H1*–H2* than initial ($\beta = -0.074$, $\chi^2(1) = 6.081$, $p = .014$) and medial positions ($\beta = -0.115$, $\chi^2(1) = 16.092$, $p < .0001$). The relation between F0 and H1*–H2* is illustrated in Figure 4. The H1*–H2* distribution at different position of different sentence type is illustrated in Figure 5. In order to visualize the effect of position controlling for F0, the residual of H1*–H2* after factoring out the effect of F0 was calculated. In order to get the residual of H1*–H2*, the measure was regressed on a model with F0 as the sole fixed effect. The coefficient of F0 was retrieved from the model summary. Residual H1*–H2* equals the raw H1*–H2* value subtracted by the product of the coefficient of F0 and the F0 value.

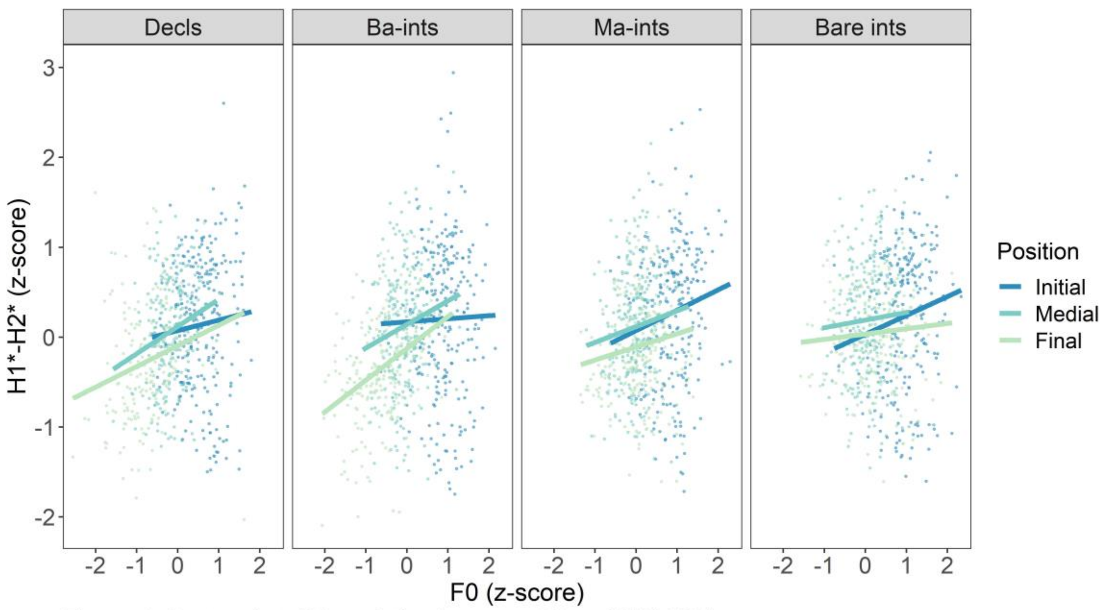

Figure 4. Scatterplot of the relation between F0 and H1*–H2*
Regression line = lm(H1*–H2* ~ F0) by sentence type and position

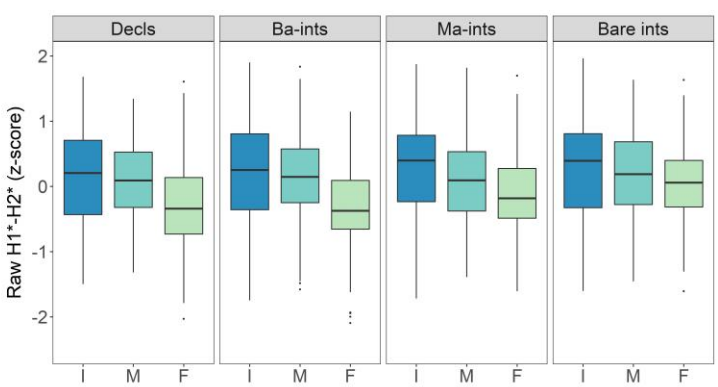

Figure 5.1. Raw H1*–H2* in different positions

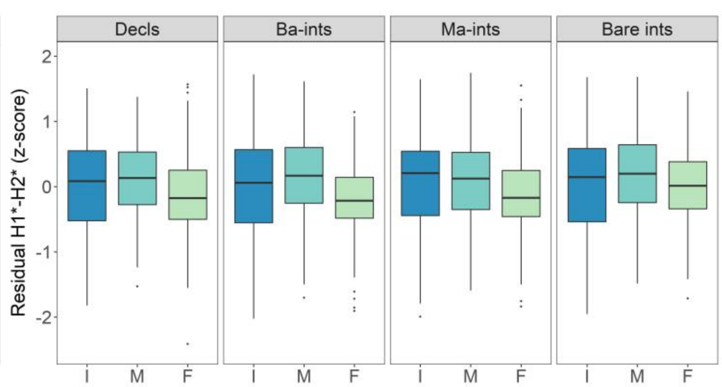

Figure 5.2. Residual H1*–H2* in different positions

4.4. HNR

The fixed and random effects of model predicting HNR are the same as the model predicting H1*–H2*. The R code for the model is (4):

(4) *lmer(HNR ~ (F0 + position) * type + (F0 + position || subject) + (F0 + position || sentence))*

Averaging across all positions and sentence types, F0 is positively correlated with HNR ($\beta = 0.248$, $\chi^2(1) = 41.114$, $p < .0001$). There was a significant interaction between F0 and sentence type ($\chi^2(3) = 12.057$, $p = .013$), and between position and type ($\chi^2(6) = 67.936$, $p < .0001$).

Using orthogonal coding to code sentence type, the difference in the position effect between declaratives vs. bare interrogatives and between "ba" interrogatives vs. "ma" interrogatives was tested within the same model. The effect of position is represented by linear (*Final - Initial*) and quadratic difference *[Medial – ½*(Initial + Final)]*. The linear effect of position was significantly different between declaratives vs. bare interrogatives ($\chi^2(1) = 19.154$, $p < .0001$), but was not significantly different between "ba" interrogatives vs. "ma" interrogatives ($\chi^2(1) = 2.266$, $p = .132$). The quadratic effect of position was significantly different between declaratives vs. bare interrogatives ($\chi^2(1) = 5.348$, $p = .021$) and between "ba" interrogatives vs. "ma" interrogatives ($\chi^2(1) = 10.119$, $p = .001$).

Since the interaction between each two factors was significant, the effect of F0 and position were tested for each sentence type separately. Since the interaction between F0 and position is not of interest of the current study, the interaction between them is omitted in the model. The R code for the model is (5):

(5) *lmer(HNR ~ F0 + position + (F0 + position || subject) + (F0 + position || sentence))*

For declaratives, controlling for utterance position, F0 was positively correlated with HNR ($\beta = 0.237$, $\chi^2(1) = 6.817$, $p = .009$). The lower the F0 was, the lower the HNR was. Controlling for F0, the HNR in final position was lower than initial ($\beta = -0.204$, $\chi^2(1) = 7.914$, $p = .005$) and medial ($\beta = -0.322$, $\chi^2(1) = 20.944$, $p < .0001$) positions.

For "ba" interrogatives, controlling for utterance position, F0 was positively related to HNR ($\beta = 0.378$, $\chi^2(1) = 26.893$, $p < .0001$). Controlling for F0, the HNR in final position was higher than initial position (β = 0.172, $\chi^2(1)$ = 8.989, p = .003), but did not differ significantly from medial position significantly (β = -0.068, $\chi^2(1)$ = 3.443, p = .064).

For "ma" interrogatives, controlling for utterance position, F0 was positively related to HNR (β = 0.219, $\chi^2(1)$ = 12.703, p = .0003). Controlling for F0, the HNR in final position was higher than initial (β = 0.289, $\chi^2(1)$ = 16.64, p < .0001) and medial (β = 0.149, $\chi^2(1)$ =8.958, p = .003) positions.

For bare interrogatives, controlling for utterance position, F0 was positively related to HNR (β = 0.245, $\chi^2(1)$ = 19.965, p < .0001). Controlling for F0, the HNR in final position was higher than initial position (β = 0.15, $\chi^2(1)$ = 10.805, p = .001), but did not differ from medial position significantly (β = -0.028, $\chi^2(1)$ = 0.61, p = .435).

The relation between F0 and HNR is illustrated in Figure 6 for each sentence type at each position separately. The distribution of HNR by position and sentence type is illustrated in Figure 7. In order to visualize the effect of position controlling for F0, the residual of HNR after factoring out the effect of F0 was calculated. In order to get the residual of HNR, the measure was regressed on a model with F0 as the sole fixed effect. The coefficient of F0 was retrieved from the model summary. Residual HNR equals the raw HNR value subtracted by the product of the coefficient of F0 and the F0 value. Table 2 and Table 3 summarize the acoustic results.

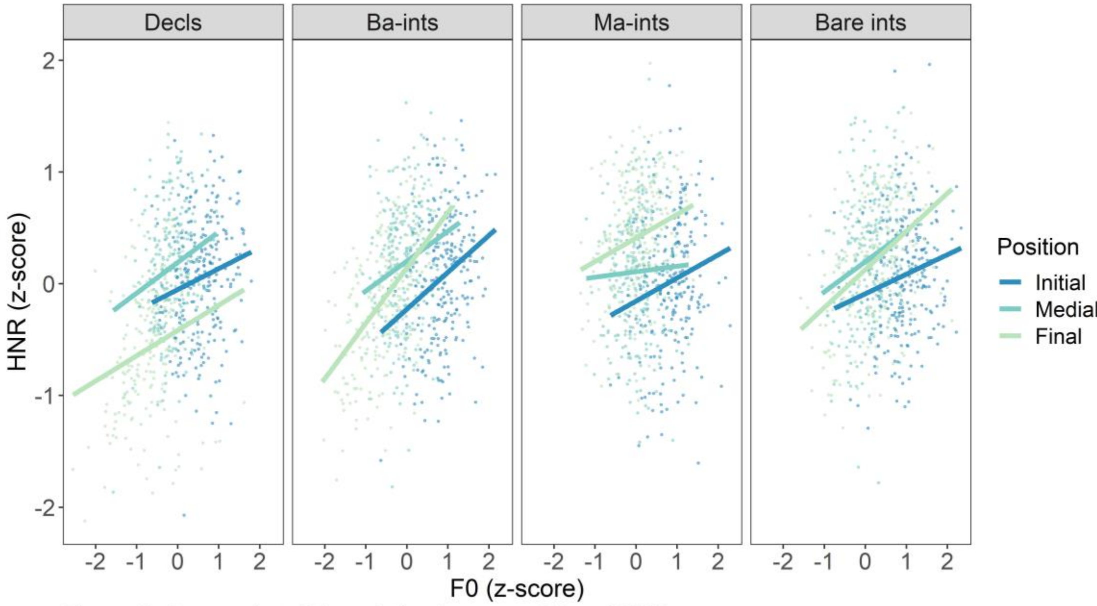

Figure 6. Scatterplot of the relation between F0 and HNR
Regression line = lm(HNR ~ F0) by sentence type and position

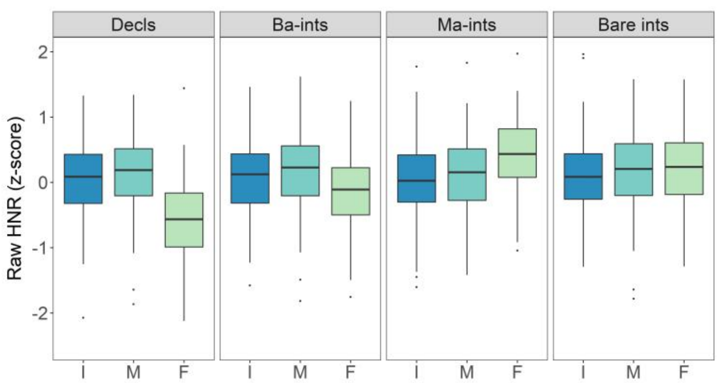

Figure 7.1. Residual HNR in different positions

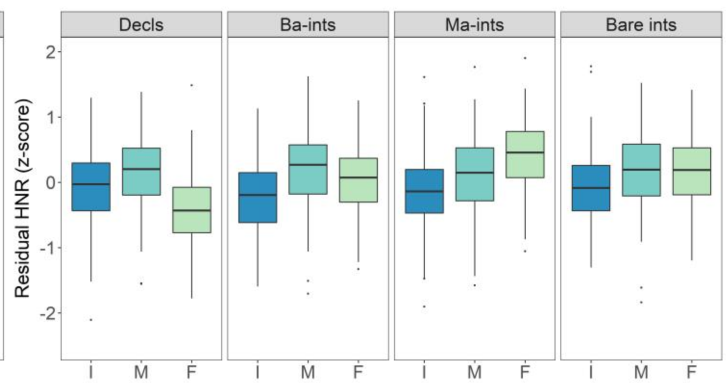

Figure 7.2. Residual HNR in different positions

Table 2. The effect of F0 and position on H1*–H2* and HNR by sentence type

	H1*–H2*		HNR	
	F0	**Position**	**F0**	**Position**
Decls			$\beta = 0.237$	F < I (Δ = -0.204) F < M (Δ = -0.322)
“Ba” ints	$\beta = 0.2$	F < I (Δ = -0.074) F < M (Δ = -0.115)	$\beta = 0.378$	F > I (Δ = 0.172) F = M (Δ = -0.068)
“Ma” ints			$\beta = 0.219$	F > I (Δ = 0.289) F > M (Δ = 0.149)
Bare ints			$\beta = 0.245$	F > I (Δ = 0.15) F = M (Δ = 0.028)

β = The slope of F0 controlling for position
Δ = The value in final position – initial or medial position, controlling for F0

Table 3. Comparison of the position effect on HNR between different sentence types

			Linear	Quadratic
Decls	**vs.**	**Bare ints**	different	different
"Ba" ints	**vs.**	**"Ma" ints**	same	different

4.5. Particles

The particles of "ma" and "ba" are unstressed and have a neutral tone, meaning they lack a pitch target. The F0 during particles is usually irregular, which makes the estimation of F0 difficult and inaccurate. Thus, the F0 of neutral tone cannot be controlled for in the model. Since H1*–H2* is computed based on F0, it cannot be calculated either. However, the HNR values of particles are illustrated in Figure 8, together with the other target positions in the utterance. The difference in HNR between "ma" vs. "ba" particles was tested using a mixed-effect model with the particle type as a fixed effect, and the intercept of subject and sentence as random effects. The HNR of "ba" particles is lower than that of "ma" particles (β = -0.932, $\chi^2(1)$ = 65.234, p < .0001).

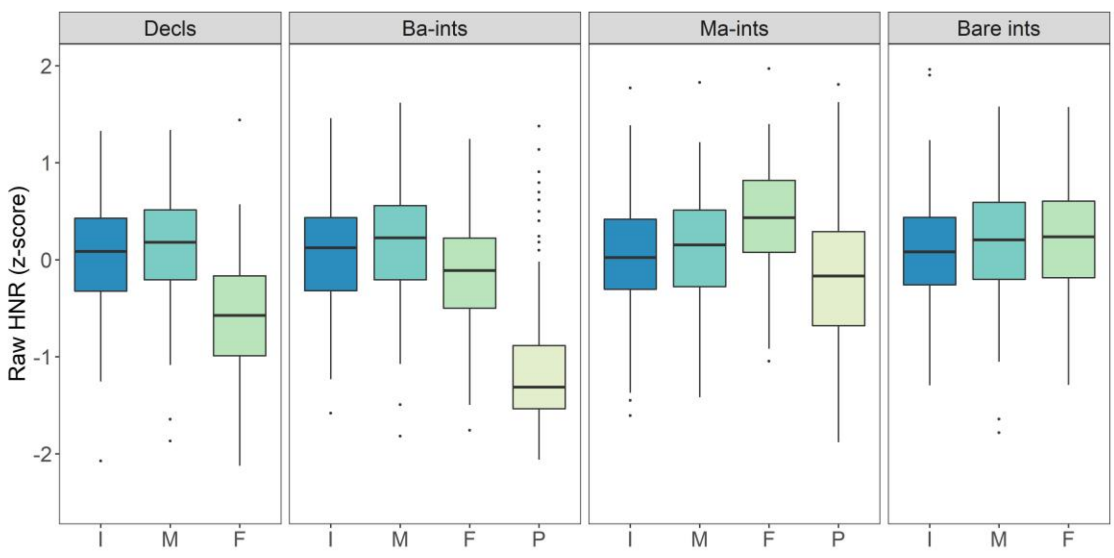

Figure 8. HNR in different positions (particle included) (P: Particle)

5. Discussion

5.1. Effects of F0 and utterance position

The first research question was: when controlling for utterance position and sentence type, what is the relation between F0 and the degree of creakiness? Overall, the results demonstrate that, in declaratives and interrogatives, lower F0 values are associated with a higher degree of vocal fold constriction (as indexed by lower H1*–H2*), and a noisier quality (as indexed by lower HNR). Together, this suggests that lower F0 is associated with creak, in accordance with Kuang (2017a).

The second research question was: when controlling for F0, is the final position of Mandarin utterances still creakier (more constricted and aperiodic) than non-final positions? The results show that utterance position has an independent effect on both the degree of constriction (denoted by H1*–H2*) and the periodicity (denoted by HNR) of the production. But final position is not always creakier than non-final positions. For all types of sentences, factoring out the effect of F0, the final position of utterance is still more constricted than non-final positions, indicating that the speakers targeted a more constricted voicing in the final position of utterances. The relation between utterance position and periodicity will be discussed in the next section.

5.2. Effect of sentence type

The third and fourth questions were: does sentence type have an independent effect on voice quality? And if it does, is the difference caused by the semantic meaning (i.e. declaratives differ from interrogatives) or the pragmatic meaning (all four types of sentences have different position effect)? The results showed that the effect of position on H1*–H2* was unvaried among the four sentence types; while the effect of position on HNR differed between declaratives and interrogatives and also among different types of interrogatives, indicating that pragmatic meaning has independent effect on voice quality. On the one hand, for declaratives, HNR declines linearly from initial to final position, indicating that the final position in declaratives is less periodic than non-final positions. On the other hand, for “ma” interrogatives, HNR *increases* linearly from initial to medial position, indicating that as the utterance proceeds, the voice quality becomes *more* periodic. For “ba” interrogatives and bare interrogatives, HNR in final position is higher than initial position but does not differ from medial position, indicating that the final position in “ba” and bare interrogatives is more periodic than initial position while the periodicity does not differ between final and medial position. The statistical results have confirmed that the linear and quadradic effects of position are significantly different between declaratives and bare interrogatives, and the quadradic effect of position is significantly different between “ma” interrogatives and “ba” interrogatives.

Based on the comparison between declaratives and bare interrogatives, and between "ba" interrogatives and "ma" interrogatives, we propose that the positional effect on voice quality is related to the pragmatic function of the sentence. The more belief the speaker has in the positive proposition, the creakier (denoted by lower HNR) the production will become in utterance-final position compared with non-final positions. The less belief the speaker has in the positive proposition, the less creaky (denoted by higher HNR) the production will become in final position compared with non-final positions.

Speakers have a more belief towards the positive proposition in "ba" interrogatives than in "ma" interrogatives. Accordingly, the HNR in final position of "ma" interrogatives is higher than that in medial position, while the HNR in final position of "ba" interrogative is marginally, though not significantly, lower than that in medial position. In addition, the particle "ba" is much noisier than the particle "ma", implying that the final position of "ba" interrogatives is less periodic than the final position of "ma" interrogatives.

Likewise, speakers have more belief towards the positive proposition denoted in declaratives than in bare interrogatives, and the results confirm that the effect of utterance position differs between declaratives and bare interrogatives. The HNR in final position of declaratives declines linearly, while the HNR of final position in bare interrogatives is higher than the initial position and equal to medial position. Thus, as the declarative sentence proceeds, the voice quality becomes less periodic, while in bare interrogatives, the voice quality becomes more periodic from the beginning to the middle, and stays consistent from the middle to the end.

"Ba" interrogatives were compared with "ma" interrogatives while declaratives were compared with bare interrogatives because the former two both end with a particle, while the latter two do not have a particle at the end. Therefore, the effect of position is comparable within each pair. But the statistical results are inconclusive in determining whether the effect of position is different between declaratives and "ba" interrogatives, because the final position of declaratives is in fact final position of the utterance, while the final position measured in "ba" interrogative refers to the word immediately before the particle. Thus, the reason why the final position of "ba" interrogatives does not differ from the medial position could be that it has not reached the absolute final position yet. And the fact that particle "ba" has a much lower HNR than the word before “ba” indicates that the speaker aims for a creaky voice quality at the end of “ba” interrogatives. Likewise, we cannot determine whether the effect of position differs between “ma” interrogatives and bare interrogatives based on the statistical results. Although the final position is more periodic in “ma” interrogatives than in bare interrogatives, this can be due to the fact that the final position measured in “ma” interrogative is not the absolute final position. The periodicity could further lower in the absolute final position. In addition, the raw HNR of “ma” particle is lower than the raw HNR of bare interrogatives, indicating that the speakers may aim for a less periodic end for “ma” interrogatives than for bare interrogatives.

Future work should compare declaratives and “ba” interrogatives, as well as “ma” interrogatives and bare interrogatives, which would require controlling for and measuring the voice quality of the segment immediately before the final segment in declaratives and bare interrogatives. That way, the “final position” would refer to the same position across four types of sentences.

An alternative interpretation of the results is that the creak pattern within utterances can be predicted by the degree of speaker’s bias towards one of the propositions denoted by the sentences. As Table 3 shows, position has an opposite effect for declaratives (I & M more periodic than F) as compared to “ma” interrogatives (I & M less periodic than F), but has the same effect for “ba” interrogatives and bare interrogatives (I less periodic than F; M equally periodic as F). On the one hand, declaratives imply a complete bias towards the only proposition denoted by the sentences, while “ma” interrogative imply no bias towards either proposition denoted by the sentences. The opposite degree of bias predicts the opposite position effect on periodicity. On the other hand, “ba” interrogatives and bare interrogatives both imply a certain degree of bias (less than 100%) towards one of the propositions denoted by the interrogative sentences. The existence of an incomplete bias seems to unify the effect of position on periodicity in those two types of sentences. However, as mentioned above, the effect of position between declaratives vs. “ma” interrogatives and between “ba” vs. bare interrogatives are not entirely comparable, due to the differences in what counts as the “final position” of those sentences. It is unclear which interpretation is correct until we control for and measure the voice quality of the segment immediately before the final segment in declaratives and bare interrogatives.

5.3. Relation between voice quality, pragmatic meaning, and prominence

The H1*–H2* and HNR results suggest that, while all types of sentences have a more constricted target in final position, they have different targets for voicing periodicity. It seems that the more belief that the speaker has in the positive proposition, the more likely that they have a periodic final target. What is the reason for this disparity? We propose that the difference in utterance prominence among those four types of sentences causes such a disparity.

Prominence is defined as words "standing out from their environment" (Streefkerk, 1997). Researchers studying the acoustic correlates of prominence have examined the difference between stressed vs. unstressed syllables and focused vs. non-focused forms. Compared with unstressed syllables, German stressed syllables are produced with more regular voicing (measured by *elastic determinism ratio* by Lancia et al., 2016), and have more energy in the higher-frequency (500–4000 Hz) bands (Sluijter & van Heuven, 1996 for Dutch). English stressed syllables are also produced with more constrictions in vocal folds (represented by lower H1*–H2* value and a longer glottal closed phase within each pulse cycle) than unstressed syllables (Garellek, 2014 (for word-initial stressed vowels only); Mooshammer, 2010). In terms of focused/pitch accented forms, some studies found that they are associated with higher F0, higher spectral intensity (perceived as loudness), and longer syllable duration (Breen et al., 2010; Terken, 1991; Vainio et al., 2010; Wang & Narayanan, 2007). Campbell and Beckman (1997) did not find consistent differences in duration between focused and non-focused forms, but confirmed that focused forms have more energy in higher frequency bands (ERB above 60). As for voice quality, Vainio et al. (2010) found that, Finnish words have less glottal closure under focus (measured by Normalized Amplitude Quotient). They claimed that focused forms are breathier than non-focused forms. Huang, Athanasopoulou, and Vogel (2018) found that, in Mandarin, the HNR value is higher in focused forms than non-focused forms of the high-level Tone 1 (55), indicating that the production of focused forms is more periodic. In summary, prominent syllable/words are usually associated with one or several of the following features: higher F0, higher intensity (especially in higher frequency), longer duration, more periodic voicing, and more/less constricted vocal folds (more constricted in stressed syllable; less constricted in focused forms).

In the current study, among the four types of utterances, the final position of “ma” interrogatives and bare interrogatives appears to be more prominent than the final position of declarative and “ba” interrogative. The final position of “ma” and bare interrogatives has a higher F0 (see Figure 2), higher voicing intensity (represented by Strength of Excitation (SoE), see Figure 9), and higher periodicity (represented by HNR, see Figure 8) than declaratives and “ba” interrogatives.

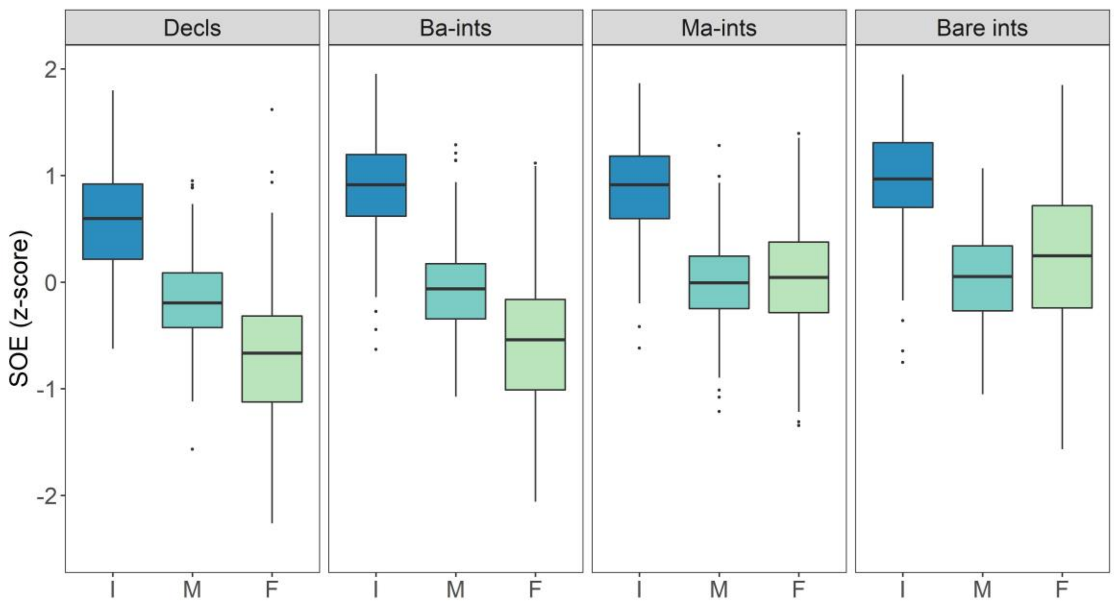

Figure 9. SoE (intensity) of different utterance positions

This begs the question why “ma” and bare interrogatives have a more prominent final position than declaratives and “ba” interrogatives. We argue that this lies in the difference in the pragmatic function among them. As mentioned above, speakers have absolute belief towards the proposition denoted by declaratives and a bias towards the positive proposition denoted by “ba” interrogatives. They are neutral towards the propositions denoted by “ma” interrogatives and have a bias towards the negative proposition denoted by bare interrogatives. It is likely that if the speaker has less belief in the positive proposition denoted by the sentence, they will emphasize the sentence in the articulation in order to draw the addressee’s attention to their doubt. The importance of emphasizing the doubt can be reflected by the expected answer from the addressee as well. For declaratives, there is no expectation on the addressee's response because there is no question asked. For "ba" interrogatives, there is a strong bias towards an affirmative answer from the interlocutor since "ba" interrogatives are inherently confirmation questions. For "ma" interrogatives, the addressee is expected to respond either "yes" or "no" since "ma" interrogatives are usually used in neutral context (Li & Thompson, 1989). For bare interrogatives, the expected answer is neither "yes" nor "no". Instead, the speaker is expressing doubt about and surprise for the positive proposition, and the interlocutor is expected to address this surprise and provide an explanation. Thus, the information load in the response is heavy for the addressee of bare interrogatives. As a result, the final position of "ma" interrogatives and bare interrogatives needs to be more prominent than declaratives and "ba" interrogatives in order to raise the addressee's attention to prepare for a more complicated response.

5.4. Type of creak

The F0 and HNR patterns differ among the four sentence types, which implies that the type of creak at utterance-final position differ among sentence type accordingly. In declaratives and "ba" interrogatives, the final position has lower F0, lower H1*–H2*, and lower HNR compared to non-final positions, indicating that the final position is more constricted and less periodic, and accompanied by low pitch. Thus, the creaky voice in the final position of declaratives and "ba" interrogatives is consistent with prototypical creak. In "ma" interrogatives and bare interrogatives, the final position has a higher F0, lower H1*–H2*, and higher HNR compared to non-final positions, indicating that the final position is more constricted but also more periodic, and is accompanied by high pitch. The creaky voice in the final position of "ma" interrogatives and bare interrogatives is consistent with the definition of a relatively tense voice (Keating et al., 2015).

6. Conclusion

The current study investigated the sources of creak in Mandarin utterances. The experiments specifically tested whether utterance-final position is creakier than non-final positions controlling for F0, and whether the voice quality pattern across positions is the same between declaratives and interrogatives and among different types of interrogatives. Three types of interrogatives, which differ in pragmatic meaning, were tested: “ba” interrogatives, “ma” interrogatives, and bare interrogatives. The results show that the effect of position is different between declaratives and interrogatives, and is not the same for all interrogatives. Controlling for F0, the final position of declaratives is still more aperiodic than non-final positions. In contrast, controlling for F0, the final position of “ma” interrogatives become more periodic than non-final positions. “Ba” interrogatives, which are used when the speaker has strong belief to the proposition, are creakier in final position than “ma” interrogatives, which are used when the speaker has neutral belief to the proposition. This study concludes that, for all types of sentences attested, the voice quality in final position differs from non-final positions controlling for F0. However, such differences between final position and non-final positions differ by sentence type. The more belief the speaker has in the positive proposition, the creakier the final position is compared with non-final positions. Mandarin speakers employ different voice qualities to differentiate the pragmatic function of sentences. Declaratives and “ba” interrogatives are characterized by prototypical creak in the final position; while “ma” and bare interrogatives are characterized by relatively tense voice in the final position.

References

Audacity computer software. (2018). Retrieved July 31, 2019, from https://audacityteam.org/

Belotel-Grenié, A., & Grenié, M. (1994). Phonation types analysis in Standard Chinese. *Proceedings of Third International Conference on Spoken Language Processing* (pp. 343-346). Retrieved from https://www.isca-speech.org/archive/icslp_1994/i94_0343.html

Belotel-Grenié, A., & Grenié, M. (2004). The creaky voice phonation and the organisation of Chinese discourse. *Proceedings of International symposium on tonal aspects of languages: With emphasis on tone languages* (pp. 5-8). Retrieved from https://www.isca-speech.org/archive_open/tal2004/tal4_005.html

Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics, 30*(2), 163-191. doi: 10.1006/jpho.2001.0155

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and cognitive processes, 25*(7-9), 1044-1098. doi: 10.1080/01690965.2010.504378

Campbell, N., & Beckman, M. (1997). Stress, prominence, and spectral tilt. *Proceedings of Intonation: Theory, models and applications* (pp. 67-70). Retrieved from https://www.isca-speech.org/archive_open/int_97/inta_067.html

Chai, Y. (2019). The source of creak in Mandarin utterances. *Proceedings of ICPhS 2019.* Retrieved from https://assta.org/proceedings/ICPhS2019/

Chasaide, A. N., & Gobl, C. (2004). Voice quality and f0 in prosody: towards a holistic account. *Proceedings of Speech Prosody 2004, International Conference* (pp. 189-196). Retrieved from https://www.isca-speech.org/archive/sp2004/sp04_189.html

Davison, D. S. (1991). An acoustic study of so-called creaky voice in Tianjin Mandarin. *UCLA Working Papers in Phonetics, 78*, 50-57. Retrieved from https://escholarship.org/content/qt8j00q68c/qt8j00q68c.pdf#page=54 de Moraes, J. A. (1999). F0 declination in Brazilian Portuguese in read and spontaneous speech. *Proceedings of ICPhS99* (pp. 2323-2326). Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/papers/p14_2323.pdf

DiCanio, C. T. (2009). The phonetics of register in Takhian Thong Chong. *Journal of the International Phonetic Association*, *39*(2), 162-188. doi: 10.1017/S0025100309003879

Fuchs, S., Petrone, C., Rochet-Capellan, A., Reichel, U. D., & Koenig, L. L. (2015). Assessing respiratory contributions to f0 declination in German across varying speech tasks and respiratory demands. *Journal of Phonetics*, *52*, 35-45. doi: 10.1016/j.wocn.2015.04.002

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of phonetics*, *24*(4), 423-444. doi: 10.1006/jpho.1996.0023

Epstein, M. A. (2003). Voice quality and prosody in English. *Proceedings of 15th ICPhS* (pp. 2405-2408). Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/p15_2405.html

Esposito, C. M. (2003). The effects of f0 and position-in-utterance on phonation in Santa Ana del Valle Zapotec. *Proceedings of 15th ICPhS* (pp. 2181-2184). Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/p15_2181.html

Garellek, M. (2013). *Production and perception of glottal stops* (Doctoral dissertation). Available from ProQuest Dissertations & Theses database. Retrieved from https://escholarship.org/uc/item/7zk830cm

Garellek, M. (2014). Voice quality strengthening and glottalization. *Journal of Phonetics*, *45*, 106-113. https://doi.org/10.1016/j.wocn.2014.04.001

Garellek, M. (2019). The phonetics of voice. *Handbook of Phonetics. New York: Routledge*.

Garellek, M., & Keating, P. (2011). The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association*, *41*(2), 185-205. doi: 10.1017/S0025100311000193

Huang, Y., Athanasopoulou, A., & Vogel, I. (2018). The Effect of Focus on Creaky Phonation in Mandarin Chinese Tones. *University of Pennsylvania Working Papers in Linguistics*, *24*(1), 12. Retrieved from https://repository.upenn.edu/pwpl/vol24/iss1/12

Keating, P. A., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. *Proceedings of ICPhS 2015.* Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0821.pdf

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, *87*(2), 820-857. doi: 10.1121/1.398894

Kreiman, J. (1982). Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics*, *10*(2), 163-175.

Kuang, J. (2017a). Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice. *The Journal of the Acoustical Society of America*, *142*(3), 1693-1706. doi: 10.1121/1.5003649

Kuang, J. (2017b). Creaky Voice as a Function of Tonal Categories and Prosodic Boundaries. *Proceedings of Interspeech* (pp. 3216-3220). Retrieved from https://www.isca-speech.org/archive/Interspeech_2017/pdfs/1578.PDF

Lancia, L., Voigt, D., & Krasovitskiy, G. (2016). Characterization of laryngealization as irregular vocal fold vibration and interaction with prosodic prominence. *Journal of Phonetics*, *54*, 80-97. doi: 10.1016/j.wocn.2015.08.001

Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. Oakland, CA: University of California Press.

Liing, W. (2014). *How to ask questions in Mandarin Chinese* New York: CUNY Academic Works.

Lin, M. (2004). On production and perception of boundary tone in Chinese intonation. *Proceedings of International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages* (pp. 125-130). Retrieved from https://www.isca-speech.org/archive_open/tal2004/tal4_125.html

Local, J. K., Wells, W. H., & Sebba, M. (1985). Phonology for conversation: Phonetic aspects of turn delimitation in London Jamaican. *Journal of Pragmatics*, *9*(2-3), 309-330. doi: 10.1016/0378-2166(85)90029-3

Maeda, S. (1976). *A characterization of American English intonation* (Doctoral dissertation). MIT.

McEnery, A., & Xiao, Z. (2004). The Lancaster Corpus of Mandarin Chinese: A corpus for monolingual and contrastive language study. *Religion*, *17*, 3-4.

Mooshammer, C. (2010). Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in German. *The Journal of the Acoustical Society of America*, *127*(2), 1047-1058. doi: 10.1121/1.3277160

Pan, H. H. (2017). Glottalization of Taiwan Min checked tones. *Journal of the International Phonetic Association*, *47*(1), 37-63. doi: 10.1017/S0025100316000281

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, *162*(1-2), 8-13. doi: 10.1016/j.jneumeth.2006.11.017

Pompino-Marschall, B., & Żygis, M. (2010). Glottal marking of vowel-initial words in German. *ZAS Papers in Linguistics*, *52*, 1-17. Retrieved from http://www.zas.gwz-berlin.de/1036.html

Redi, L., & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, *29*(4), 407-429. doi: 10.1006/jpho.2001.0145

Ruan, L. (2004). *A study of the intonation of the interrogative sentence in Mandarin Chinese* (Unpublished master's thesis). Beijing Language and Culture University, Beijing, China.

Sharoff, S. (2006). Creating general-purpose corpora using automated search engine queries. *WaCky*, 63-98. doi: 10.1.1.136.5005

Shen, X. N. S. (1990). *the prosody of Mandarin Chinese* (Vol. 118). Oakland, CA: University of California Press.

Shih, C. (1997). Declination in mandarin. *Proceedings ofIntonation: Theory, Models and Applications 1997* (pp. 293-296). Retrieved from https://www.isca-speech.org/archive_open/int_97/inta_293.html

Shue, Y.-L., P. Keating , C. Vicenik, K. Yu (2009) VoiceSauce: A program for voice analysis, *Proceedings of 16th ICPhS*, (pp. 1846-1849).

Sluijter, A. M., & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical society of America*, *100*(4), 2471-2485. doi: 10.1121/1.417955

Smith, C. L. (2002). Prosodic finality and sentence type in French. *Language and Speech*, *45*(2), 141-178. doi.org/10.1177/00238309020450020301

Stevens, K. N. (2000). *Acoustic phonetics* (Vol. 30). Cambridge, MA: MIT press.

Streefkerk, B. M. (1997). Acoustical correlates of prominence: A design for research. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* (Vol. 21, pp. 131-142).

Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *The Journal of the Acoustical Society of America*, *142*(4), 2469-2482. doi: 10.1121/1.5008854

Terken, J. (1991). Fundamental frequency and perceived prominence of accented syllables. *The Journal of the Acoustical Society of America*, *89*(4), 1768-1776. doi: 10.1121/1.401019

Vainio, M., Airas, M., Järvikivi, J., & Alku, P. (2010). Laryngeal voice quality in the expression of focus. *Proceedings of Interspeech 2010* (pp. 921-924). Retrieved from https://www.isca-speech.org/archive/interspeech_2010/i10_0921.html

Wang, D., & Narayanan, S. (2007). An acoustic measure for word prominence in spontaneous speech. *IEEE transactions on audio, speech, and language processing*, *15*(2), 690-701. doi: 10.1109/TASL.2006.881703

Yu, K. M., & Lam, H. W. (2014). The role of creaky voice in Cantonese tonal perception. *The Journal of the Acoustical Society of America*, *136*(3), 1320-1333. doi: 10.1121/1.4887462

Yuan, J., & Liberman, M. (2014). F0 declination in English and Mandarin broadcast news speech. *Speech Communication*, *65*, 67-74. doi: 10.1016/j.specom.2014.06.001

Yuan, J., Shih, C., & Kochanski, G. (2002). Comparison of declarative and interrogative intonation in Chinese. Retrieved from https://www.ling.upenn.edu/~jiahong/publications/c23.pdf

Zeng, X. L., Martin, P., & Boulakia, G. (2004). Tones and intonation in declarative and interrogative sentences in Mandarin. *Proceedings of International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages* (pp. 235-238). Retrieved from https://www.isca-speech.org/archive_open/tal2004/tal4_235.html

Zhang, H. (2016). Boundary effects on allophonic creaky voice: A case study of Mandarin lexical tones. *Proceedings of Tonal Aspects of Languages 2016*, 94-98.

Zheng, X. (2006). Voice quality variation with tone and focus in mandarin. *Proceedings of Tonal Aspects of Languages* (pp. 132-136). Retrieved from https://www.isca-speech.org/archive/tal_2006/tal6_132.html

Zhuang, X., & Hasegawa-Johnson, M. (2008, May). Towards interpretation of creakiness in switchboard. *Proceedings of Speech Prosody 2008* (pp. 37-40). Retrieved from http://isle.illinois.edu/sprosig/sp2008/