

The Neutral Tone in Question Intonation in Mandarin

Fang Liu¹, Yi Xu²

¹ Department of Linguistics, The University of Chicago, IL, USA

² Department of Phonetics and Linguistics, University College London, UK

liufang@uchicago.edu, yi.xu@ucl.ac.uk

Abstract

This study investigates how the neutral tone, when preceded by different full tones under different focus conditions, behaves in question intonation in Mandarin. Results indicate that 1) the preceding full/neutral tone largely determines the local F₀ trajectory of the neutral tone, but the latter also gradually converges over the course of several neutral tone syllables, 2) post-focus lowering, which is caused by the effect of focus, occurs in both neutral-tone-ending and High-tone-ending sentences, with the interrogative intonation in questions realized as an upward shift starting from the focused word, and 3) sentence-final neutral tone has a falling contour even in questions, thus contrasting with sentence-final High tone, which has a rising contour in questions.

Index Terms: neutral tone, focus, sentence type

1. Introduction

There are four full lexical tones and a neutral tone in Mandarin. On a five-point scale [1], the High tone in citation form is transcribed as /55/, the Rising tone /35/, the Low tone /214/, and the Falling tone /51/. The pitch of the neutral tone is determined by its preceding tone: high after the Low tone, and relatively low after the other tones [2,3]. Nevertheless, in investigating the carryover and anticipatory effects of full tones on a sequence of 1-3 neutral tones, Chen & Xu [4] concluded that the neutral tone has a static mid target, which is implemented with weak articulatory strength.

Although the basic characteristics of the four full tones in connected speech do not deviate much from their citation forms, there have been reports that intonation and sentence environment do modify their register and contour characteristics: 1) The overall pitch values of the four full tones in sentence-final position are higher in questions than in statements [5-11]. 2) The High tone falls moderately at the final position of a declarative sentence [12,13], but rises in an interrogative sentence [6,9,11]. 3) The steepness of the Rising and Falling tone of the final syllable differs in questions and statements [8], with the Rising tone steepened [10] and widened [7,9,11] and the Falling tone flattened [10,11] and narrowed [7] in pitch range in questions. 4) According to [14], the sentence-final Low tone is realized as falling in statements but falling-rising in questions. However, in [7] and [11], the pitch realization of the final Low tone is falling-rising in both sentence types, except that the rising ending is higher in questions than in statements. The data in [6] further reveal that only focused Low is realized as falling-rising statement-finally, and non-focused Low is falling in a statement-final position. On the other hand, the question-final Low tone is realized as falling-rising no matter whether it is focused or not.

The issue of tone-intonation interaction becomes even trickier when the neutral tone is involved. If indeed targetless as traditionally viewed [2,3] or with a weakly implemented mid target [4], the F₀ of the neutral tone at the final position of

a question should fully or at least largely reflect the interrogative intonation. Previous research, however, have produced mixed results. In Chao's impressionistic account [2], for a "Low + High + Rising + Neutral + Neutral + Neutral" sequence, "nǐ /21/ shuō /55/ shén /35/ me /2/ lái /1/ zhe /1/? [What did you say?]", the last three neutral tones are realized with a low pitch if the question is newly-uttered, but with a high pitch ("nǐ /21/ shuō /55/ shén /35/ me /5/ lái /5/ zhe /5/?") if the same question is asked again. Qi [15], however, stated that the neutral-tone question-particle "ma" has a high pitch /5/ no matter what the preceding full tone is. Furthermore, with a sequence of three neutral tones plus "ma" at the end of a question, their pitch can be transcribed as /4 4 4 5/ regardless of what the preceding full tone is. Yet, upon observing that the question-final neutral-tone particle "ma" has a falling contour after the High and Falling tone, a rising contour after the Low tone, and a falling/level/rising contour after the Rising tone, Shen [9:48] claimed that the tonal value of the sentence-final interrogative particles is "the algebraic sum of the F₀ value of the preceding tone and the sentence intonation", and that they "always end on a high key". In the Pan-Mandarin ToBI system [16], the sentence-final particle "ma" can carry two possible boundary tones: H% in questions and L% in statements. However, Lee [17:122] proposed that "the ma-particle does not stand as an independent prosodic unit that could bear the terminal rise" in marked questions.

In summary, previous research on Mandarin has not reached a consensus on the pitch pattern of the neutral tone in interrogative sentences, and the picture presented is far from complete. There was no explicit indication of how the preceding tone and the sentence intonation interact to generate the F₀ contour of a question-final neutral tone in the algebraic sum hypothesis in [9]. The limited examples in [16] and [17] also failed to present a full picture of the influence of different preceding tones on the neutral tone in different intonations. Chen & Xu [4] have neither examined the neutral tone at the sentence-final position nor in questions. Therefore, this study attempts to address the following unresolved issues regarding the neutral tone in question intonation in Mandarin: (1) Which of the two has a stronger influence on the F₀ trajectory of the neutral tone: the preceding tone or sentence type? (2) Does the neutral tone have different targets in different sentence types? (3) What is the effect of focus on the neutral tone syllables in statements and questions? (4) Is the neutral tone more effective than a full tone in manifesting the statement/question distinction?

2. Method

2.1. Materials

Tā mǎi māma/yéye/nǎinai/mèimei men de le ma
[He bought mothers'/grandpas'/grandmas'/sisters']

Tā mǎi māma/yéye/nǎinai/mèimei men de māomī
[He bought mothers'/grandpas'/grandmas'/sisters' kitten]

The above two sentence frames are employed in the experiment, onto which the following three conditions are imposed to create 32 combinations.

- **Focus: nonadjacent vs. adjacent.** The focus is either (on *mǎi*) separated from the sequence of 5 neutral tones (or 3 neutral tones + 2 High tones) by a full tone, or (on *mā/yé/nǎi/mèi*) immediately adjacent to them.
- **Preceding tone: High, Rising, Low, or Falling.** The sequence of 5 neutral tones (or 3 neutral tones + 2 High tones) is preceded by each of the four full tones.
- **Sentence type: statement vs. question.** Each utterance with the same components was produced with two alternate sentence types.

Each utterance was repeated five times by each subject, resulting in a total of 1280 sentences. The intended focus and sentence type were elicited by proper leading sentences.

2.2. Subjects

Eight native speakers of Mandarin, 4 females and 4 males, served as subjects. They were either students at Yale University or residents in New Haven, Connecticut, who were born and raised in the city of Beijing where Mandarin is the vernacular. They had no self-reported speech and hearing disorders and their ages ranged from 23 to 34.

2.3. Procedure

Recording was done in a sound-treated booth at Haskins Laboratories, New Haven, Connecticut. A JavaScript program displayed the target sentences (and the leading sentences) one at a time on a computer screen in random order. The recorded utterances were digitized at 44.1 kHz sampling rate and 16-bit amplitude resolution, and were later re-sampled at 22.05 kHz. For visual inspection and graphic analysis, a custom-written script for the Praat program [18] computed time-normalized F_0 contours of the sentences by getting the same number of evenly spaced F_0 points from each syllable.

3. Results

To see how sentence type is affected by lexical tone and focus, time-normalized F_0 contours of the statements and questions with focus on *mǎi* and *mā/yé/nǎi/mèi* are displayed in Fig. 1 and 2, respectively, each averaged across 40 repetitions by eight speakers.

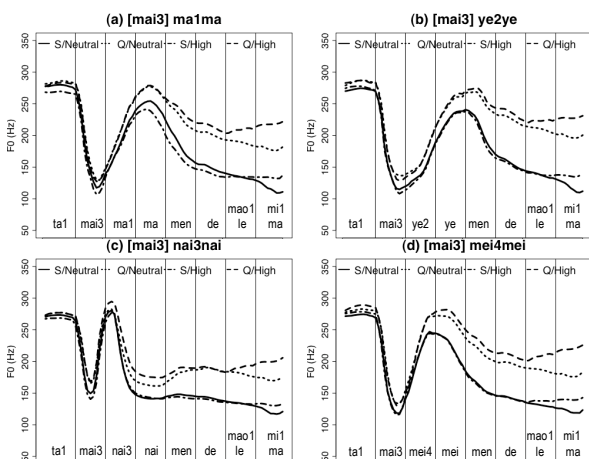


Fig. 1: F_0 contours of statements/questions with focus on “*mǎi*”. In the legend, “S/Neutral” refers to a statement ending with 5 neutral tones, “Q/High” a question with 2 High tones, and so on. The title

“*[mai3] ma1ma*” indicates that focus is on “*mǎi*” and the 3rd and 4th syllables of the sentence are “*māma*”.

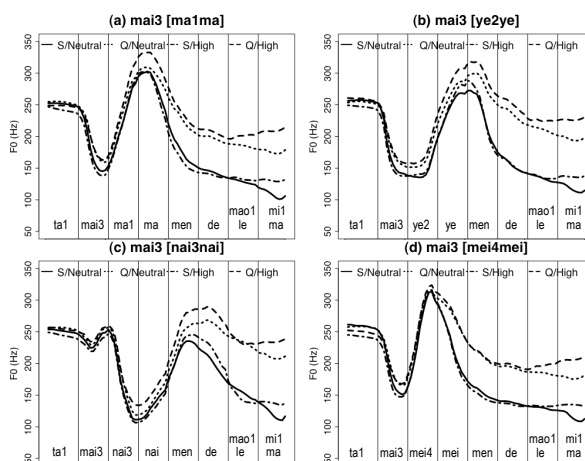


Fig. 2: F_0 contours of statements/questions with focus on “*mā/yé/nǎi/mèi*”. The title “*mai3 [ma1ma]*” indicates that focus is on “*māma*”, the 3rd and 4th syllables of the sentence.

As can be seen, regardless of the tone of the third (*mā/yé/nǎi/mèi*) and last two syllables (neutral vs. High), in each graph, questions differ from statements with an increasingly higher F_0 starting from the focused syllable (*mǎi* in Fig. 1 and *mā/yé/nǎi/mèi* in Fig. 2) till the end of the sentence. Furthermore, this difference is larger in the final portion of the sentence when the last two tones are High than when they are neutral. Regarding the effect of focus, comparing Fig. 1a-d with Fig. 2a-d, we can see that focused *mǎi* not only lowers its own F_0 but also raises the F_0 of the preceding syllable *tā*, whereas focused *mā/yé/nǎi/mèi* expands its own pitch range as well as that of the following neutral-tone syllable *mā/yé/nǎi/mèi*.

In Fig. 1c and 2c, due to the tone sandhi rule (Low + Low > Rising + Low) in Mandarin [3], the second syllable *mǎi* (Low) changes into *mái* (Rising) before the third syllable *nǎi* (Low). Because focus is on *mái* in Fig. 1c but on *nǎi* in Fig. 2c, the rising contour of *mái* is more fully realized in the former than in the latter. Furthermore, the F_0 trajectories of the rest of the syllables differ dramatically in the two figures because of the different focus locations. In Fig. 1c, F_0 of the neutral-tone syllable *nai* remains low when preceded by the post-focus Low-tone syllable *nǎi*, but that of the next neutral-tone syllable *men* increases slightly, which is in turn followed by a slow decrease in the rest of the neutral-tone syllables *de le ma*. In Fig. 2c, however, following the focused Low-tone syllable *nǎi*, F_0 of the next two neutral-tone syllables *nai men* rises sharply for the most part, but starts to fall during *de* and continues to fall throughout *le ma*. This phenomenon is described as “post-L F_0 rise” in [4]. The current finding suggests that this effect is quite weak when the Low tone itself is post-focus.

To see how the neutral tone sequences (and the 3-neutral plus 2-High sequences) behave after each of the four full tones under different focus and sentence type conditions, F_0 contours of the statements and questions alternating across *māma/yéye/nǎinai/mèimei* on the third and fourth syllable with the neutral-tone and High-tone ending are displayed in Fig. 3 and 4, respectively, each averaged across 40 repetitions by eight speakers.

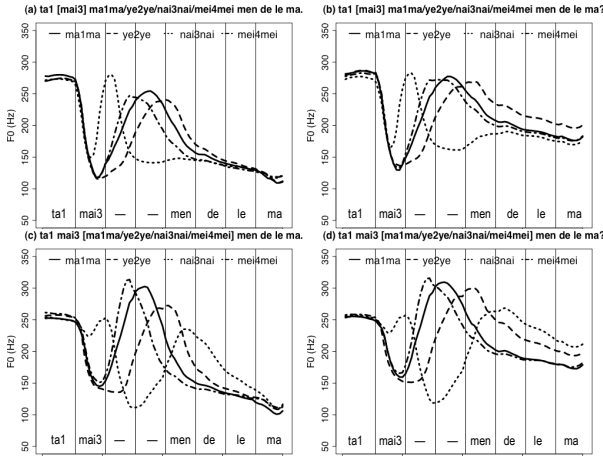


Fig. 3: F_0 contours of statements/questions alternating across “māma/yéye/nǎinai/mèimeī” on the third and fourth syllables with focus on “mǎi” (in (a) and (b)) or on “mā/yé/nǎi/mèi” (in (c) and (d)) and with the neutral tone on the final 5 syllables.

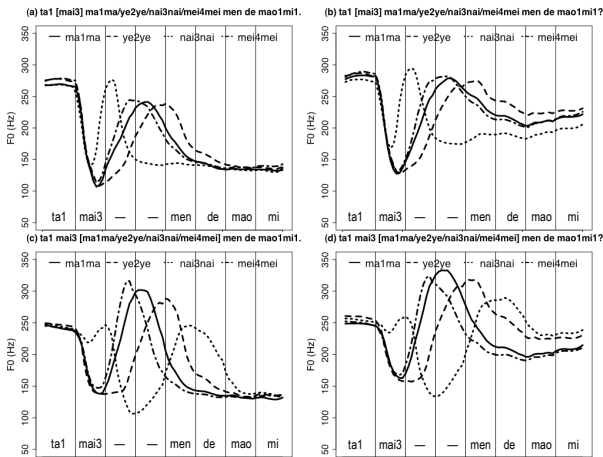


Fig. 4: F_0 contours of statements/questions alternating across “māma/yéye/nǎinai/mèimeī” on the third and fourth syllable with focus on “mǎi” (in (a) and (b)) or on “mā/yé/nǎi/mèi” (in (c) and (d)) and with the High tone on the final two syllables.

Fig. 3 shows that the magnitude of the influence of the full tone on the neutral tones depends not only on the tonal target of the full tone, but also on the focus and sentence type conditions. In Fig. 3a, where the sentences are statements with focus on *mǎi*, the effect of the full tones on the neutral tone sequence shows the descending order of Rising > High > Falling > Low, with the last two neutral tones largely converged to a falling contour and thus free from such effect. In Fig. 3c, where the sentences are also statements but the focus is on *mā/yé/nǎi/mèi*, the Low tone has the most significant effect on the trajectory of the neutral tone sequence, while the order of the effects from the other three tones (Rising > High > Falling) remains the same. Here only the last neutral tone seems to have mostly converged to a falling contour. In Fig. 3b and 3d, where the corresponding questions under different focus conditions are plotted, the last one or two neutral tones do not converge because the Rising tone (Fig. 3b) and also the Low tone (Fig. 3d) elevate the last four neutral tones further than the other tones do. However, this elevation does not change the overall decreasing trend of the final neutral tones in question intonation.

What is observed above for the neutral tone sequence also applies to the 3 neutral tones plus 2 High tones sequence in Fig. 4, except that the High-tone ending acquires a slight rising contour in questions, as opposed to its level pattern in statements.

A set of repeated measures ANOVAs were conducted using R [19] to see if the observations made on Fig. 1-4 are statistically significant. In each of the ANOVA models, mean F_0 (in semitone) of each of the eight syllables in the sentences is treated as the dependent variable, sex as the between-subjects factor, and focus, sentence type, final tone, and preceding tone as the within-subjects factors. Results of the main effects from these models are shown in Table 1.

Table 1. Results of the main effects from repeated measures ANOVAs of mean F_0 of each syllable on sex (Female vs. Male), focus (Nonadjacent: on “mǎi” vs. Adjacent: on “mā/yé/nǎi/mèi”), sentence type (Question vs. Statement), final tone (High vs. Neutral), and preceding tone (High, Rising, Low, or Falling). The p -values less than 0.05 are bolded.

		sex	focus	sentence type	final tone	preceding tone
tā	F	15.30	7.73	2.58	3.75	1.85
	p	0.0079	0.0320	0.1596	0.1008	0.1738
		F>M	A<N	Q>S	H<N	R>F>H>L
mǎi	F	17.54	0.00	8.05	14.16	27.16
	p	0.0058	0.9996	0.0297	0.0094	< 0.0001
		F>M	A>N	Q>S	H<N	L>F>H>R
mā yé nǎi mèi	F	13.53	15.66	25.61	5.32	138.94
	p	0.0104	0.0075	0.0023	0.0605	< 0.0001
		F>M	A>N	Q>S	H>N	F>L>H>R
ma ye nai mei	F	14.67	13.14	35.74	2.89	108.67
	p	0.0087	0.0110	0.0010	0.1400	< 0.0001
		F>M	A>N	Q>S	H>N	H>F>R>L
men	F	20.64	69.31	51.81	1.87	58.08
	p	0.0039	0.0002	0.0004	0.2200	< 0.0001
		F>M	A>N	Q>S	H>N	R>H>L>F
de	F	27.91	56.07	79.82	1.81	51.60
	p	0.0019	0.0003	0.0001	0.2274	< 0.0001
		F>M	A>N	Q>S	H>N	L>R>H>F
le māo	F	33.16	11.73	116.73	7.19	35.51
	p	0.0012	0.0141	< 0.0001	0.0365	< 0.0001
		F>M	A>N	Q>S	H>N	L>R>H>F
ma mī	F	38.42	0.41	153.67	37.74	14.18
	p	0.0008	0.5456	< 0.0001	0.0009	< 0.0001
		F>M	A>N	Q>S	H>N	R>L>F>H

The main effects shown in Table 1 can be summarized as follows. 1) Female speakers generally have higher F_0 than male speakers. 2) The effect of focus varies with its location: (i) When it is on *mǎi*, which is nonadjacent to the following neutral tone sequence, focus has an anticipatory raising effect on the mean F_0 of the preceding syllable *tā*. (ii) When focus is on *mā/yé/nǎi/mèi*, which is adjacent to the following neutral tone sequence, it raises the mean F_0 of both the focused and the post-focus syllables. 3) Starting from the focused item, the F_0 of questions becomes higher than that of statements. The difference in F_0 height between the two is increasingly larger as the sentences approach the end. This is demonstrated by the more and more significant effect of sentence type on the mean F_0 of the eight consecutive syllables in the sentences. 4) There is no obvious anticipatory effect of the tone of the final two syllables (High vs. neutral) on the mean F_0 of the preceding syllables. Rather, the effect is strictly local: sentences ending with High tones have significantly higher final mean F_0 than those ending with neutral tones.

5) According to the tone sandhi rule (L + L > R + L) in Mandarin, *māi* (L) becomes *mái* (R) before *nāi* (L), which is why the mean F₀ of *māi* is the highest before the Low tone (when compared to the other three). Following *māi*, the mean F₀ of the four full tones shows the order of *mèi* > *nāi* > *mā* > *yé*. This is due to the articulatory constraint that determines how fast F₀ can change even with full effort [20]. Following the four full tones, there are four corresponding neutral tone syllables whose F₀ trajectory seems to be a continuation of its preceding full tone, thus having mean F₀ in the order of *ma* > *mei* > *ye* > *nai*. Following the ending point of *māma/yéye/nāinai/mèimei*, the F₀ contour of the neutral tone syllable *men* is mostly falling after *māma/yéye/mèimei*, but mostly rising after *nāinai* (especially when *nāinai* is focused). This makes its mean F₀ the highest after *yéye*, but much lower after *māma/nāinai/mèimei*. The F₀ contour of *de* is falling no matter what the preceding full tone is, but the “post-L F₀ rise” of the focused *nāi* on *nai men* raises the pitch range of their following *de* so that this *de* has a higher mean F₀ than those preceded by the other three full tones. Although most cases of the last two syllables *le ma/māomī* converge to a falling/level contour in statements, only those preceded by *māma men de/mèimei men de* converge in questions. In addition, *le ma/māomī* are in a relatively higher pitch range when following focused/unfocused *yéye men de* and focused *nāinai men de*.

4. Discussion and Conclusions

Conventionally considered to be “targetless”, the neutral tone at the question-final position has been assumed to be fully free to carry the rising intonation and therefore should be high in pitch [15]. However, the present study suggests that the neutral tone is not necessarily any better than a full tone in manifesting question intonation, because it does not exhibit a high pitch in a question, but rather behaves differently according to the tone and focus conditions of the preceding syllables. Furthermore, post-focus lowering occurs in the neutral-tone-ending sentences in both statements and questions, just as in the High-tone-ending sentences.

In conclusion, this study shows that 1) the F₀ trajectory of the neutral tone is mainly determined by the preceding tone, which itself is heavily influenced by the focus and intonation conditions of the sentence; 2) when the F₀ contours of the neutral tone in different conditions converge over time, they appear to be slightly falling in both statements and questions. Nonetheless, the fall is steeper in statements than in questions; 3) the effects of focus and sentence type are exerted on the neutral tone via global pitch controls. In this experiment, the neutral tone sequence occurs in the post-focus region and at the sentence boundary. Thus, the post-focus pitch range suppression due to focus and the nonlinear post-focus pitch increase due to question intonation [6] combine to form the pitch pattern of the neutral tone sequence (and the neutral tone + High tone sequence); 4) the F₀ of the sentence-final High tones remains level in statements, but becomes rising in questions. This is in contrast to the F₀ of the final neutral tones, which keeps falling not only in statements but also in questions. These patterns indicate that, in Mandarin, sentence type is an independent intonational function whose manifestation has to be achieved through the articulatory implementation of local tonal pitch target specified by the lexical tones, including the neutral tone.

5. References

- [1] Chao, Y. R., “A system of ‘tone letters’”, *Le Maître Phonétique*, 45: 24–27, 1930.
- [2] Chao, Y. R., “Tone and Intonation in Chinese”, *Bulletin of the Institute of History and Philology*, 4: 121-134, 1933.
- [3] Chao, Y. R., *A Grammar of Spoken Chinese*, University of California Press, Berkeley, CA, 1968.
- [4] Chen, Y. and Xu, Y., “Production of weak elements in speech -- Evidence from F₀ patterns of neutral tone in standard Chinese”, *Phonetica*, 63: 47-75, 2006.
- [5] Ho, A. T., “Intonation variations in a Mandarin sentence for three expressions: interrogative, exclamatory, and declarative”, *Phonetica*, 34: 446-456, 1977.
- [6] Liu, F. and Xu, Y., “Parallel Encoding of Focus and Interrogative Meaning in Mandarin Intonation”, *Phonetica*, 62: 70-87, 2005.
- [7] Ni, J. and Kawai, H., “Pitch targets anchor Chinese tone and intonation patterns”, *Proceedings of Speech Prosody 2004*, Nara, Japan, 95-98, 2004.
- [8] Rumjancev, M. K., “Ton i intonacija v sovremennom kitajskom jazyke [Tone and Intonation in Modern Chinese]”, *Izdatel'stvo Moskovskogo Universiteta*, Moscow, 1972. Reviewed by Lyovin, A. V., *Journal of Chinese Linguistics*, 6: 120-168, 1978.
- [9] Shen, X.-N. S., *The Prosody of Mandarin Chinese*. University of California Press, Berkeley, 1990.
- [10] Yuan, J., *Intonation in Mandarin Chinese: Acoustics, Perception, and Computational Modeling*, Ph.D. dissertation, Cornell University, 2004.
- [11] Zeng, X.-L., Martin, P., and Boulakia, G., “Tones and intonation in declarative and interrogative sentences in Mandarin”, *Proceedings of International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages*, Beijing, 235-238, 2004.
- [12] Ho, A. T., “Mandarin tones in relation to sentence intonation and grammatical structure”, *Journal of Chinese Linguistics*, 4: 1-13, 1976.
- [13] Ho, A. T., “The acoustic variation of Mandarin tones”, *Phonetica*, 33: 353-367, 1976.
- [14] Lin, M., “On production and perception of boundary tone in Chinese intonation”, *Proceedings of International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages*, Beijing, 125-129, 2004.
- [15] Qi, S., “Hanyu de zidiao, tingdun yu yudiao de jiaohu guanxi [The interaction among lexical tones, pause, and intonation in Chinese]”, *Zhongguo yuwen*, 1956(10): 10-13, 1956.
- [16] Peng, S.-H., Chan, M. K.M., Tseng, C.-Y., Huang, T., Lee, O. J., and Beckman, M. E., “Towards a Pan-Mandarin system for prosodic transcription”, in Sun-Ah Jun (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, Oxford University Press, Oxford, UK, 230-270, 2005.
- [17] Lee, O. J., *The prosody of questions in Beijing Mandarin*, Ph.D. dissertation, Ohio State University, 2005.
- [18] Boersma, P. and Weenink, D., *Praat: doing phonetics by computer*, <http://www.praat.org/>, 2006.
- [19] R Development Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, URL <http://www.R-project.org>, 2005.
- [20] Xu, Y. and Sun, X., “Maximum speed of pitch change and how it may relate to speech”, *Journal of the Acoustical Society of America*, 111: 1399-1413, 2002.