Mechanisms of Question Intonation in Mandarin

Jiahong Yuan

Department of Linguistics, University of Pennsylvania
Philadelphia, PA 19104, USA
jiahong@ling.upenn.edu

Abstract. This study investigates mechanisms of question intonation in Mandarin Chinese. Three mechanisms of question intonation have been proposed: an overall higher phrase curve, higher strengths of sentence final tones, and a tone-dependent mechanism that flattens the falling slope of the final falling tone and steepens the rising slope of the final rising tone. The phrase curve and strength mechanisms were revealed by a computational modeling study and verified by the acoustic analyses as well as the perception experiments. The tone-dependent mechanism was suggested by a result from the perceptual study: question intonation is easier to identify if the sentence-final tone is falling whereas it is harder to identify if the sentence-final tone is rising, and was revealed by the acoustic analyses on the final Tone2 and Tone4.

Keywords: question intonation, perception, acoustics.

1 Introduction

This study investigates mechanisms of question intonation in Mandarin Chinese. In terms of surface F0, the difference between question and statement intonation in Mandarin Chinese is quite diverse, which can be illustrated by Figure 1.
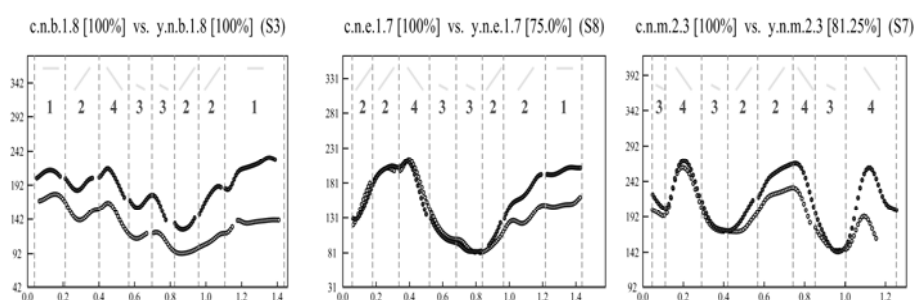

Fig. 1. Diverse patterns of difference between the question curve (dark circles) and the statement curve (open circles).

In general, the difference can be realized as the following: 1. The question curve is higher than the statement curve on the whole sentence (the left pair in Figure 1). 2. The question curve diverges from the statement curve after a point (the middle pair in Figure 1). 3. The question curve is higher than the statement curve except at some portions or points (the right pair in Figure 1).

The difference between question and statement intonation has attracted much attention in Chinese intonation study. De Francis claims that the whole pitch level of the interrogative is higher than that of the declarative [1]. Disagreeing with De Francis, Tsao argues that the whole pitch level has no difference between the two intonation types and interrogative intonation in Chinese is 'a matter of stress' [2]. Gårding models Chinese intonation with 'grids', which qualitatively mark a time-varying pitch range. Lexical tones then fit into that range [3, 4]. In Gårding's model, the two intonation types have different grids. Shen J. proposes that the top line and the base line of a pitch contour are independent in the prosodic system of Chinese [5, 6]. For interrogative intonation the top line falls gradually whereas the base line undulates slightly and ends at a much higher point (compared to declarative intonation). Shen X. investigates the difference between the two intonation types by comparing their pitch values at four points: starting point, highest peak, lowest trough, and ending point [7]. Her conclusion is that interrogative intonation begins at a register higher than declarative, although it may end with either a high or low key. In Pan-Mandarin ToBI, question intonation is mainly associated with a high boundary tone in the intonational tones tier [8].

We studied the difference between question and statement intonation in Chinese with Stem-ML [9], an intonation description language combined with an algorithm for translating tags into quantitative prosody. Our study found that the 'diverse' difference between question and statement intonation in Mandarin Chinese can be accounted for by two consistent mechanisms: an overall higher phrase curve for the question intonation, and higher strength values of sentence final tones for the question intonation. It also suggested that the phrase curves of the two intonation types tend to be parallel and boundary tones are not necessary for modeling the difference between the two intonation types in Mandarin Chinese [10].

These results raised several interesting issues and questions: Firstly, they seem inconsistent with some previous models of Chinese intonation. For example, in Gårding's model the two intonation types have grids with different directions and in Pan-Mandarin ToBI a high boundary tone has been used to transcribe question intonation. Secondly, if question intonation is realized by strengthening the final syllables then what is the difference between 'plain' interrogative intonation and interrogative intonation whose final syllable(s) are focused (given the fact that a focused syllable must also be strengthened in speech)? And what is the difference between 'plain' question intonation and declarative intonation that has a focus at its end? Thirdly and most importantly, do actual production and perception data support these results?

Perception and acoustic studies were then conducted in light of these issues and questions. A systematic corpus was created for both of these facets of the study. The corpus consisted of 130 sentences, which were minimal pairs contrasting on intonation type (statement, question), presence of a focus or not, focus position

(sentence initial, middle, end), tone of the focused syllable (tone1, tone2, tone3, tone4), and tone of the last syllable (tone1, tone2, tone3, tone4). For example:

1. Li3bai4wu3 Luo2Yan4 yao4 mai3 mao1. ← statement, no focus
   Friday   Luo2Yan4 will   buy   cat
   "Luo2Yan4 (Luo2 is a last name and Yan4 is a first name) will buy a cat Friday."
2. Li3bai4wu3 Luo2yan4 yao4 mai3 mao1? ← question, no focus
   Friday   Luo2yan3 will   buy   cat
   "Luo2Yan4 will buy a cat Friday?" (It is a question with surprise, dubiousness, etc.)
3. Li3bai4wu3 Luo2Yan4 yao4 mai3 **mao1**. ← statement, a tone1 focus at the end
   Friday   Luo2Yan4 will   buy   **cat**
   "(not a goat, not a deer,) LuoYan will buy a **cat** Friday."

Eight native Mandarin speakers, four male and four female, took part in the recording. Two perception experiments were conducted on the 1040 utterances recorded in the database. One was for identifying intonation type and the other for identifying focus. Sixteen listeners, 8 female and 8 male, participated in the perception experiments. The listeners are also native Mandarin speakers.

Results of the perception experiments have been reported in [11, 12]. In section 2 I summarize these results, and argue that the perception experiments found evidence for the strength mechanism of question intonation, as well as evidence for a third mechanism that was not found in our previous Stem-ML modeling study. In section 3, I report the results of acoustic analyses, which show further evidence for the strength and the phrase curve mechanisms, and discuss what the third mechanism is. Finally, in section 4, I present the conclusions.

2 Evidence from Perception Experiments

As reported in [11, 12], the perception experiments found that: 1. Statement intonation is easier to identify than question intonation; 2. The tone of the last syllable does not affect the identification of statement intonation; 3. The tone of the last syllable does affect the identification of question intonation: First, question intonation is easier to identify on a sentence ending with Tone4 than those ending with the other tones; second, identification of some speakers' question intonation is very difficult if the sentence ends with Tone2; 4. A focus at the end of a sentence makes statement intonation more difficult to identify; 5. A focus at the middle or the end of a sentence makes question intonation easier to identify; 6. A focus at the middle of a sentence is easier to identify than a focus at the beginning or at the end, no matter what intonation type the sentence has; 7. A focus at the end of a sentence is more difficult to identify than a focus at the beginning for a statement but not for a question; 8. The tone of a focused syllable does not affect focus identification under Statement, even if the focus is at the end of a sentence; 9. A focus on Tone2 is more difficult to identify than that on Tone4 if the focus is at the middle of a question sentence and a focus on Tone1 is more difficult to identify than that on Tone3 if it is at the end of a question sentence. Some of the results were also found in Liu and Xu's study on a different dataset [13].

These results reveal four interesting asymmetries: statement and question intonation identification; effects of the final tone2 and tone4 on question intonation identification; effects of the final focus on statement and question intonation identification; and effects of intonation type on focus identification.

The asymmetry of statement and question intonation identification manifests in two ways: First, statement intonation is easier to identify than question intonation, suggest by both a higher mean identification ratio and a smaller variation; second, the tone of the last syllable does not affect statement intonation identification but it does affect question intonation identification. The intonation identification test was a forced choice test: the listeners must identify the intonation type of each utterance as either a statement or a question. That question intonation identification was less accurate means that many question intonation utterances were identified as statements. This suggests that statement intonation is a default or unmarked intonation type. That is, listeners fall back to this option when there is not enough information suggesting 'question', which is also supported by the fact that the tone of the last syllable does not affect Statement identification. Question intonation is, however, a marked intonation type. It can only be identified if the listeners actually hear the 'question' features/mechanisms.

The second asymmetry revealed by the perception experiments is of the effects of the sentence-final Tone2 and Tone4 on question intonation identification: On the one hand, question intonation is easier to identify on a sentence ending with Tone4 than ending with the other tones. On the other hand, identification of some speakers' question intonation is very difficult if the sentence ends with Tone2. Tone4 is a falling tone and Tone2 is a rising tone. Therefore the asymmetry can also be stated as follows: In sentence-final position, question intonation is easier to identify on a falling tone and more difficult to identify on a rising tone. Our previous Stem-ML modeling study on Mandarin intonation showed that the difference between statement and question intonation in Mandarin Chinese can be accounted for by two mechanisms: an overall higher phrase curve for the question intonation, and higher strength values of sentence final tones for the question intonation. Neither of these gestures, however, seems to be able to explain this asymmetry. The raise of the phrase curve is a global mechanism and has nothing to do with the asymmetry of the local interaction between intonation and the final tone. If the strength mechanism accounts for the asymmetry, at the sentence final position a high strength on Tone2 should be more difficult to identify than that on Tone4. Our perception test on focus identification, however, does not support this hypothesis. The mechanism of a focus must be a high strength, but the tone of the focused syllable, according to conclusion 8 above, does not affect focus identification under Statement, even if the focus is at the end of a sentence. The inability of the strength mechanism to explain this asymmetry implies either of the following two conclusions: 1. the strength mechanism is a 'false' one; 2. there exists another mechanism that accounts for the asymmetry. The first implication, however, is not tenable, as we can see from the discussion below as well as in section 3.

The third and fourth asymmetries are related to both focus and intonation type. The third asymmetry is that a focus at the end of a sentence makes statement intonation harder to identify but makes question intonation easier to identify. And from conclusion 6 and 7 we can generalize the fourth asymmetry: Question intonation makes a focus at the end of a sentence easier to identify whereas statement intonation

does not. Both of the asymmetries are consistent with the strength mechanism of question intonation. Both question intonation and a final focus have a higher strength at the sentence final position. Therefore, presence of both in a sentence will make it easier for the listeners to identify the higher strength mechanism, which is an indicator of question intonation to the listeners in the intonation type identification test and an indicator of focus in the focus identification test. If there is a focus at the end of a statement, the higher strength of the last focused tone may be misinterpreted as a mechanism of question intonation for some listeners. Therefore more statements were identified as questions if focus was presented in final position.

   In summary, the perception experiments found evidence for the strength mechanism of question intonation. It also found an asymmetry about the effects of the sentence-final tone2 and tone4 on question intonation identification. That the strength and the phrase curve mechanisms cannot explain this asymmetry suggests there is another mechanism of question intonation.

3   Evidence from Acoustic Analyses

Comparisons of statement and question intonation were made on the overall F_0 pattern, the overall intensity and duration pattern, the intensity and duration of the final tones, and the F_0 of the final Tone2 and Tone4. Only the unfocused utterances were used for the acoustic analyses.

3.1   Overall F_0 Pattern

Since each statement and question intonation pair has the same tone sequence, we can calculate the difference of the average F_0 over the tone at each syllable pair. The results are shown in Figure 2.
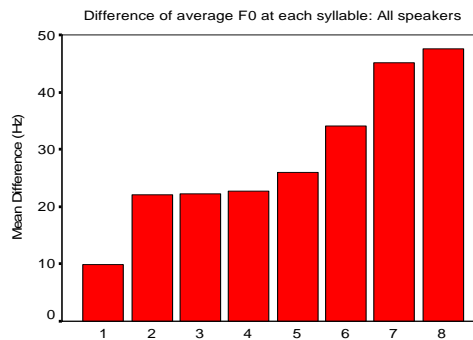

Fig. 2. Difference of the average F_0 over each syllable pair (question minus statement) for all speakers.

Clearly, the F_0 curve of question intonation is higher overall than that of statement intonation. We can also see that the difference widens toward the end of the sentence. This is consistent with the mechanisms of question intonation we found in our Stem-ML modeling study of Chinese intonation: Question intonation has a higher phrase curve and higher strengths at sentence final tones.

On the other hand, although in general the F_0 curve of question intonation is higher than that of statement intonation, Tone3 in a question may reach the same low point as in a statement at any sentence position, as we can see from Figure 1 above. To further support the conclusion that Tone3 in statement and question intonation may reach the same low point at any sentence position, Table 1 provides the counts of the occurrences where Tone3 reaches the same or almost the same low point (the difference is less than 5 Hz) at different sentence positions.

Table 1. Counts of the occurrences of the same low Tone3s in statements and questions.

Speaker	Sentence position of Tone3			
	Syllable 1	Syllable 3	Syllable 5	Syllable 7
S1	2 (11)*	2 (16)	0 (20)	0 (14)
S2	2 (11)	1 (16)	0 (20)	0 (14)
S3	4 (11)	3 (16)	0 (20)	0 (14)
S4	3 (11)	1 (16)	2 (20)	1 (14)
S5	1 (11)	1 (16)	0 (20)	2 (14)
S6	9 (11)	11 (16)	6 (20)	5 (14)
S7	0 (11)	10 (16)	7 (20)	4 (14)
S8	6 (11)	10 (16)	7 (20)	2 (14)
Total	27 (88)	39 (128)	22 (160)	14 (112)

* The numbers in the brackets are the total occurrences.

3.2 Overall Duration and Intensity Pattern

Each utterance has eight syllables. Each syllable has its own duration and intensity. Duration is the time span of the syllable and intensity is measured by the highest intensity value of the syllable. The mean duration and intensity of each syllable in statement and question intonation across all speakers are shown in Figure 3 and 4 respectively.

Figure 3 shows that the syllables in question intonation are shorter than those in statement intonation in every position except the last syllable, which is longer in question intonation. It also shows final lengthening of both statement and question intonation, because of which the final syllable is longer than the other syllables.

Figure 4 shows that question intonation has a higher intensity curve than statement intonation and that the difference between them grows toward the end of the sentence.

The duration and intensity patterns shown in Figures 3 and 4 strongly support the strength mechanism: The sentence final tones have higher strength values in question intonation.
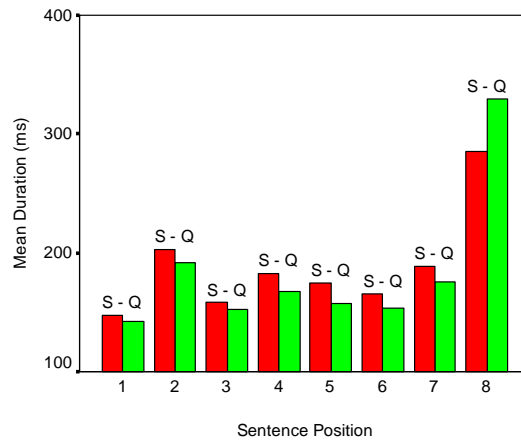
Fig. 3. Overall duration pattern of statement and question intonation (S: Statement; Q: Question).
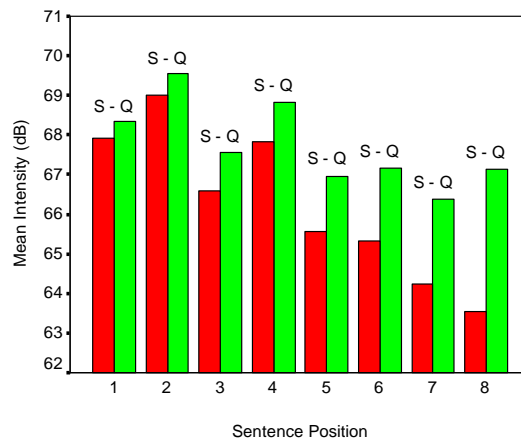

Fig. 4. Overall intensity pattern of statement and question intonation (S: Statement; Q: Question).

3.3 Duration and Intensity of the Final Tones

Figure 5 shows the duration and intensity difference between statement and question intonation for each of the sentence final tones.
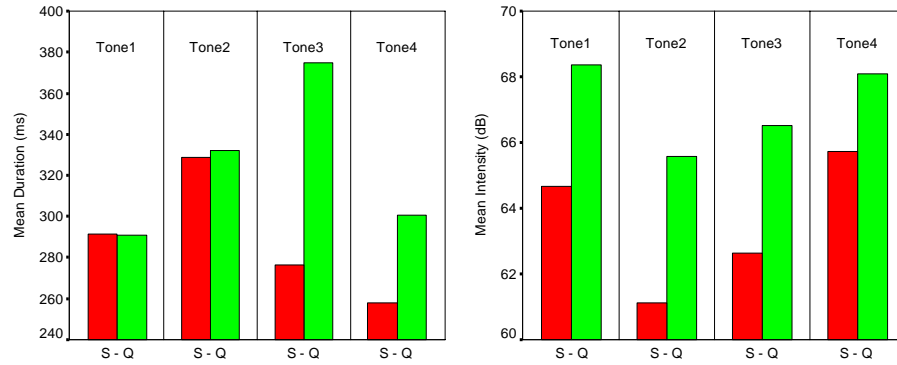
Fig. 5. Duration and intensity difference between statement and question intonation for each of the final tones (S: Statement; Q: Question).

From Figure 5 we can see that the final Tone3 and Tone4 are longer in question intonation than in statement intonation, whereas the final Tone1 and Tone2 have similar duration in the two intonation types.

In many cases the final Tone3 is only a low target in a statement. In a question, however, the rising end of Tone3 appears. Therefore, it is not surprising that the final Tone3 in question intonation is much longer than in the statement. However, why the final Tone4 is longer in question intonation whereas the final Tone1 and Tone2 are not is puzzling.

We can also see from Figure 5 that each final tone has a higher intensity in question intonation than in statement intonation. The difference between them, however, is the largest for Tone2 and the smallest for Tone4.

3.4 F_0 of the Final Tone2 and Tone4

Both the perception experiments and the duration and intensity analyses above revealed an asymmetry between the final Tone2 and Tone4. Do they also show an asymmetry in F_0? This section tries to answer this question. Two F_0 parameters are extracted and compared for the final Tone2 and Tone4: F_0 at the end and F_0 slope. The results are shown in Figures 6 and 7.

From Figures 6 and 7 we can see that both the F_0 at the end and the F_0 slope of the final Tone2 are higher in question intonation than in statement intonation. The F_0 at the end of the final Tone4 is higher in question intonation than in statement intonation whereas the F_0 slope is not different between the two intonation types.
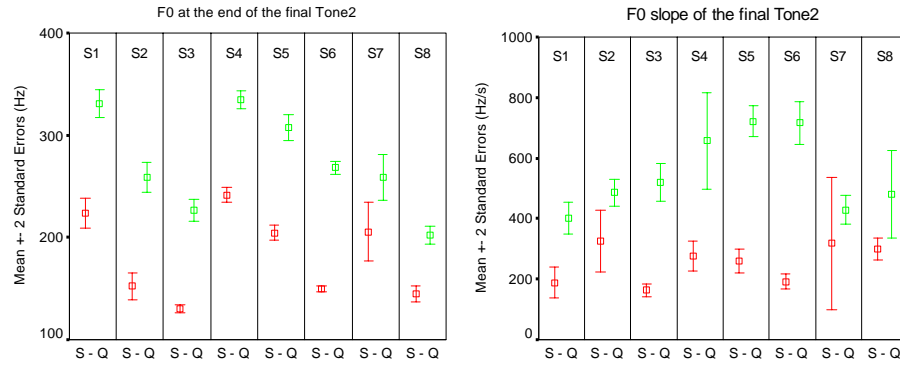
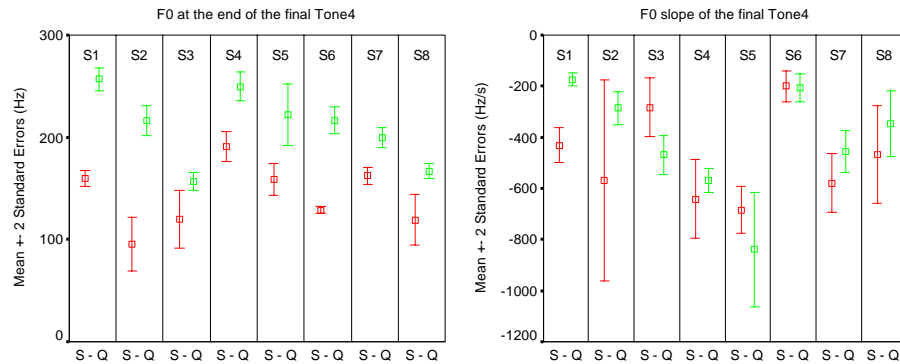**Fig. 6.** F_0 of the final Tone2 (S: Statement; Q: Question).

Fig. 7. F_0 of the final Tone4 (S: Statement; Q: Question).

3.5 Summary and Discussion

To summarize the results of the acoustic analyses, I draw the following conclusions:
1. The F_0 curve of question intonation is higher than that of statement intonation. Tone3, however, may sometimes pull the question curve down to the statement curve.
2. Question intonation has a higher intensity curve than statement intonation and the difference between them becomes greater toward the end of the utterance. 3. The final Tone3 and Tone4 are longer in question intonation than in statement intonation whereas the final Tone1 and Tone2 have similar duration in the two intonation types. 4. Each final tone has a higher intensity in question intonation than in statement intonation and the difference between them is the largest for Tone2 and the smallest for Tone4. 5. Both the F_0 at the end and the F_0 slope of the final Tone2 are higher in question intonation than in statement intonation. 6. F_0 at the end of the final Tone4 is

higher in question intonation than in statement intonation whereas the F_0 slope is not different between the two intonation types.

The perception results in section 2 suggest that a third mechanism is needed to explain the asymmetry in the effects of the sentence-final Tone2 and Tone4 on question intonation identification. I will argue that the third mechanism is a tone-dependent mechanism functioning on the final tone.

The final tone in question intonation is strengthened by the strength mechanism, which expands the F_0 range of the final tone. If there is no other mechanism functioning on the final Tone4 in question intonation, the F_0 at its end should stay low, as under focus [14]. The acoustic analyses above, however, showed that the F_0 at the end of the final Tone4 is higher in question intonation than in statement intonation. Therefore, there must be another mechanism causing the F_0 at the end of the final Tone4 in question intonation to be raised, therefore higher than in statement intonation.

From the acoustic analyses we know that the F_0 at the end and the F_0 slope of the final Tone2 are greater in question intonation than in statement intonation. I propose that the third mechanism will increase the slope of the final Tone2 in question intonation if the slope can be naturally increased. Why can it not raise the F_0 at the end? According to previous studies [14], sentence final Tone2 has a longer duration under focus (strengthened). The sentence final Tone2 in question intonation (also strengthened), therefore, must be longer than in statement intonation if the third mechanism of question intonation does not affect the duration of the final Tone2. However, the acoustic analyses above show the contrary: The duration of the final Tone2 in question intonation is not longer than that found in statement intonation. If we assume that the third mechanism of question intonation raises the F_0 at the end of the final Tone2, it will be very difficult to explain why the final Tone2 in question intonation is not longer than in statement intonation. There is no such difficulty if we assume that changing the rising slope is the mechanism. If the slope is steeper, it will take less time to finish a *span*, which is, from the speakers' point of view, a pitch range that is large enough to indicate question intonation and also natural.

From the discussion above, we can see that the third mechanism of question intonation is different from the phrase curve mechanism and the strength mechanism in the following ways: First, it is a strictly local mechanism that functions only on the last tone; second, it is tone dependent; it flattens the falling slope of the final Tone4 and steepens the rising slope of the final Tone2. The intuition of the mechanism is simple: If there is a falling tone at the end of question intonation, the falling tone goes down more slowly; if there is a rising tone at the end of question intonation, the rising tone goes up more quickly. The idea that the realization of intonation type is sensitive to tonal identity was first developed in Shih (1988) [15]. Most of the Chinese intonation models in the literature do not, however, capture this aspect of tone and intonation interaction.

The tone-dependent mechanism may conflict with the strength mechanism on the final Tone2. This mechanism requires that the final Tone2 in question intonation go up more quickly or as soon as possible. The strength mechanism, however, requires that the beginning low part of the final Tone2 in question intonation be lengthened, or not go up soon [14]. This probably explains why question intonation is more difficult to realize if there is a Tone2 at the end of a sentence.

I have paid little attention to the F_0 contours of the final Tone1 and Tone3, partly because they are difficult to study: In an utterance, Tone1 can be either a target or a high-level contour and Tone3 can have either a rising end or not. Although the third mechanism is tone dependent, whether it functions on different versions of Tone1 (having a level contour or not) or Tone3 (having a rising end or not) in the same way or in different ways is neither clear nor a trivial question, and remains a question for further research.

4 Conclusions

Results from the perception experiments demonstrate that statement intonation is easier to identify than question intonation, and, while the tone of the last syllable does not affect statement intonation identification, it does affect question intonation identification. The intonation identification test was a forced choice test: listeners must identify the intonation type of each utterance as either statement or question intonation. That question intonation identification was less accurate means that many question intonation utterances were identified as statement intonation. This suggests that statement intonation is a default or unmarked intonation type; listeners fall back on this option when there is not enough information suggesting 'question.' The conclusion is also supported by the fact that the tone of the last syllable does not affect statement intonation identification. Question intonation, however, is a marked intonation type. It can only be identified if listeners actually hear the 'question' features/mechanisms.

Three mechanisms of question intonation have been proposed: an overall higher phrase curve, higher strengths of sentence final tones, and a tone-dependent mechanism that flattens the falling slope of the final falling tone and steepens the rising slope of the final rising tone. The phrase curve and strength mechanisms were revealed by the computational modeling study and verified by the acoustic analysis as well as the perceptual study: 1. Overall, the F_0 and intensity of question intonation are higher than statement intonation (phrase curve mechanism). 2. The F_0 and intensity difference between question intonation and statement intonation becomes higher toward the end of the sentence (strength mechanism). 3. The syllables in question intonation are shorter than those in statement intonation in every position except the last syllable, which is longer in question intonation (strength mechanism). 4. Focus at the end of a sentence makes statement intonation harder to identify but makes question intonation easier to identify. Question intonation makes focus at the end of a sentence easier to identify whereas statement intonation does not (strength mechanism). The third mechanism, a tone-dependent mechanism, was suggested by a result from the perceptual study: Question intonation is easier to identify if the sentence-final tone is falling whereas it is harder to identify if the sentence-final tone is rising. Neither the phrase curve mechanism nor the strength mechanism can explain this result. The phrase curve mechanism is global and tone-independent and the strength mechanism is partially global and also tone-independent. The asymmetry of the effects of the final Tone2 and Tone4, however, is local and tone-dependent. Acoustic analyses on the final Tone2 and Tone4 suggested that the tone-dependent

mechanism of question intonation flattens the final falling tone and steepens the final rising tone.

The tone-dependent mechanism may conflict with the strength mechanism on the final Tone2: This mechanism requires that the final Tone2 in question intonation go up more quickly or as soon as possible. The strength mechanism, however, requires that the beginning low part of the final Tone2 in question intonation be lengthened, or not go up soon. This likely explains why question intonation is more difficult to realize as well as identify if there is a Tone2 at the end of a sentence.

References

1. DeFrancis, J. F.: *Beginning Chinese*. New heaven: Yale University Press (1963)
2. Tsao, W.: Question in Chinese. *Journal of Chinese Language Teachers' Association* (1967) 15-26
3. Gårding, E.: Constancy and variation in Standard Chinese tonal patterns. *Lund University Working Papers 28, linguistics-phonetics* (1985) 19-51
4. Gårding, E.: Speech act and tonal pattern in Standard Chinese: constancy and variation. *Phonetica* (1987) 13-29
5. Shen, J.: Beijinghua shengdiao de yinyu he yudiao[Pitch range of tone and intonation in Beijing dialect]. in *BeijingYuyin Shiyanlu,* Lin, T.; Wang L. (ed.). Beijing: Beijing University Press (1985)
6. Shen, J.: Hanyu yudiao gouzao he yudiao leixing [Intonation structure and intonation types of Chinese]. *Fangyan* (1994) 221-228
7. Shen, X.: *The Prosody of Mandarin Chinese,* University of California Press (1989)
8. Peng, S., Chan, M., Tseng, C., Huang, T., Lee, O., 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, U.K. (2005) 230-270
9. Kochanski, G.P., Shih, C.: Prosody modeling with soft template. *Speech Communication* (2003) 311-352
10. Yuan, J., Shih, C., Kochanski, G.P.: Comparison of declarative and interrogative intonation in Chinese. In *Proceedings of Speech Prosody 2002*. Aix-en-Provence, France (2002) 711-714
11. Yuan, J., Shih, C.: Confusability of Chinese Intonation. In *Proceedings of Speech Prosody 2004*. Nara, Japan (2004) 131-134
12. Yuan, J.: Perception of Mandarin Intonation. *Proceedings of ISCSLP 2004*. Hong Kong (2004)
13. Liu F., Xu Y.: Parallel Encoding of Focus and Interrogative Meaning in Mandarin Intonation. *Phonetica* (2005) 70-87
14. Yuan, J.: *Intonation in Mandarin Chinese: Acoustics,Perception, and Computational Modeling*, Ph.D. Dissertation, Cornell University, Ithaca (2004)
15. Shih, C.: Tone and Intonation in Mandarin. *Working Papers of the Cornell Phonetics Laboratory* (1988) 83-109