

Perceptual Studies in Chinese Intonation Using Foreign Listeners

Jiālù Zhāng

At the Department of Linguistics and Phonetics, Lund University, studies of tone and intonation in Standard Chinese have been carried out by using both natural and synthetic speech. The declination as an intrinsic feature of speech pitch has been investigated by Cohen et al. 1982. In this paper, synthetic speech is used to investigate how the baseline and the topline of the pitch contour influence the perception of sentence intonation and how focus is expressed in a sentence.

Using the LPC speech synthesis program, we can change any part of the pitch contour but preserve the spectrum parameters. One tone contour could also be changed into another one by changing only the fundamental frequency. However, we discovered that intensity is another important factor in intonation, not only for whispered, but also for normal speech, so care must also be taken to study the interaction between fundamental frequency and intensity in intonation.

Synthesized Chinese sentences in which the pitch contour was changed were presented to a group of Swedish listeners and ABX tests were carried out.

TESTING MATERIAL

The following three test sentences designed for the study of tone and intonation in Standard Chinese were used for the perceptual experiment (the pīnyīn transcription is used):

1. Wāng Yī chōu xiāngyān. ("Wāng Yī smokes cigarettes.")
2. Sòng Yàn qùguò Dàqīng. ("Sòng Yàn went to Dàqīng.")
3. Lǐ Xiǎobǎo xiě jiǎngyǎn gǎo. ("Lǐ Xiǎobǎo writes lecture notes.")

(Tone marks: ˉ stands for high level tone, ˋ for falling tone, and ˊ for low dipping tone.)

All these sentences were processed on a microcomputer with the LPC program, and the pitch contours were abstracted. The pitch contour was changed on the screen of a graphic terminal, and using the new contour the test sentences were synthesized. Sentences 1 and 2 were used for ABX tests and sentence 3 was evaluated only by the author.

All pitch contours of sentences 1 and 2, which were presented to listeners in synthesized sentences, are shown in Figures 1 and 2. Every pitch contour of a test sentence is accompanied by a test number followed by the highest frequency of the ordinate and a target mark (S, Q or F). Test sentences marked S have a target sentence which is the answer to a yes-no question with focus on the word shì "yes" (such as Shì, Wāng Yī chōu xiāngyān "Yes, Wāng Yī smokes cigarettes"). The focussed word shì was omitted. Q stands for a yes-no question (segmentally the same as a statement), and F for a sentence with focus at some place. In some cases, where the change in pitch was difficult to see on the contours, frequency values are indicated above the curves.

TESTING METHODS

1. ABX Tests

In order to investigate how changing only the pitch contour influences intonation, a special testing crew which consisted of 16 persons (10 males and 6 females, aged from 20 to 61 years, most of them native speakers of Swedish) was organized. The listeners were teachers or students of phonetics, who knew no Chinese. The instruction words: "You should make a decision based only on the intonation for every test item" were given to all listeners before the tests. Thus no syntactic or semantic information besides the pitch

contour was utilized by these foreign listeners. In this experiment, A is the reference sentence - a neutral (focus free statement) sentence -, B is the target sentence, and X is the tested sentence. Stimuli A and B were presented in random sequence and X followed them. A pause of about 10 seconds was made between the tests to allow the listeners to make their responses. The listeners were seated in an acoustically treated classroom and listened via a loudspeaker.

2. Subjective evaluation

The third test sentence (Lǐ Xiǎobǎo xiě jiǎngyǎn gǎo) was evaluated only by the author (who is a native speaker of Standard Chinese).

RESULTS AND DISCUSSION

The listeners' responses for test sentences 1 and 2 under different testing conditions (see Figures 1 and 2) are listed in Tables 1 and 2. In these tables, A or B stands for the listener's response to the test sentence X, and the percentage shows the relative number of responses made for the target sentence, i.e. B. If the target response percentage is greater than 50%, the effect of the pitch change is considered as positive and denoted "yes", and if the percentage is less than or equal to 50%, the effect is considered negative and denoted "no". These responses were based only on the pitch contour, because none of the listeners had any knowledge of Chinese. In my opinion, using foreign listeners is a good way to exclude the influence of syntactic and semantic factors in speech, while preserving the intonation.

The results of this experiment can be summarized as follows:

1. From tests No. 1, 7, 8 and 9 in Table 2, it can be seen that the final part of a pitch contour (a word or a syllable), and especially its baseline, is very important for a question sentence. Test No. 1 in Table 2, which shows that a perfect question sentence can be formed by shifting only the baseline of the last syllable qǐng upwards, while

preserving the topline, is convincing evidence. This has a similar effect as test No. 14 in Table 2, where the whole pitch contour was changed into the shape of a question sentence.

2. However, if the baseline of each syllable in a sentence is shifted upwards while preserving the topline (test No. 2 in Table 2), then there is no effect. The same phenomenon can be observed in test No. 1 in Table 1, in which the topline of xiāngyān was shifted upwards. Comparing test No. 2 to test No. 9 in Table 2, we find that the response percentage increases rapidly when both the topline and the baseline are shifted upwards. So it can be concluded that high pitch level at the end of the sentence and both topline and baseline with positive slope (maybe with a steeper slope in the baseline) are distinctive characteristics of question sentences.

3. By shifting down the baseline only (test No. 4 in Table 2) or the topline only (test No. 15 in Table 2), we cannot make a neutral sentence into a focus left statement sentence. For forming a focus left statement sentence it is necessary to shift both the topline and the baseline downwards, giving them a negative slope, much steeper for the topline than for the baseline (see test No.6 in Table 1).

4. Focus is formed by expanding the pitch range of the focussed lexical unit and compressing the pitch in other parts of the sentence, especially in the part of the sentence which follows the focus. For all tones with the exception of the dipping (third) tone, shifting the topline upwards is sufficient for forming a focus. For the dipping tone, however, the baseline has to be shifted downwards and stay low for a long time (sometimes with a rising end). A perfect focus was formed on the subject of sentence 1 (Wāng Yī) by shifting the topline of this word upwards and shifting it down for other words (see test No. 8 in Table 1). If only the topline of the subject in sentence 2 (Sòng Yàn) is shifted upwards and the rest is preserved, this does not form a focus (test No. 11 in Table 2).

5. Besides the fundamental frequency, intensity may be another important factor in the perception of intonation. All

efforts to change a neutral sentence in which all syllables have the dipping tone (such as sentence 3) into a question sentence by shifting the baseline and the topline upwards, as was done for sentences No. 1 and 2, were defeated. Segmenting the speech sounds in a synthesized sentence (test sentence 3), I found that even the baseline of the final syllable gǎo could be shifted greatly upwards without it being heard, because of the masking effect of the preceding part with higher intensity and lower frequency.

6. By changing the tone contour only, we can change any tone into any other tone, with the exception of the dipping tone at the end of a sentence.

ACKNOWLEDGEMENTS

I wish to express my gratitude to Professor Eva Gårding, who heads the project, and to Engineers Bengt Mandersson and Karl-Gustav Jönsson for their help in finishing this work.

REFERENCES

Cohen, A., R. Collier and J. 't Hart. 1982. "Declination: Construct or intrinsic feature of speech pitch?". Phonetica 39, 254-73.

Table 1. Listeners' responses for ABX tests.

Test sentence: Wāng Yī chōu xiāngyān.

Date: 25 March 1983.

Test No.	L i s t e n e r																Correct	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	%	
1	A	B	A	A	A	A	A	A	A	B	A	A	A	A	A	B	19	No
2	A	B	B	A	A	B	B	A	A	A	A	A	A	A	A	B	32	No
3	A	B	A	B	A	A	A	A	A	A	B	A	B	A	A	B	32	No
4	B	B	B	B	B	B	A	A	A	B	A	A	B	A	B	B	62	Yes
5	B	A	B	B	B	B	B	B	B	B	A	B	B	B	A	B	81	Yes
6	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	94	Yes
7	B	A	B	A	B	A	B	B	B	B	B	B	B	A	B	B	75	Yes
8	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	B	94	Yes
9	B	A	B	B	B	A	B	B	B	B	B	B	B	B	A	B	81	Yes

Table 2. Listeners' responses for ABX tests.

Test sentence: Sōng Yàn qūguò Dāqīng.

Date: 25 March 1983.

Test No.	L i s t e n e r																Correct	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	%	
1	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	A	87	Yes
2	B	A	A	A	B	B	A	A	A	B	B	B	B	B	B	B	62	Yes
3	B	B	B	B	B	A	B	A	B	A	B	B	B	A	B	A	62	Yes
4	A	A	B	A	B	A	A	A	A	A	A	A	A	A	A	A	13	No
5	B	A	A	B	B	B	A	B	A	B	A	B	B	B	B	A	62	Yes
6	A	B	B	B	B	B	B	B	A	B	B	B	B	A	B	B	81	Yes
7	A	B	B	B	B	A	A	B	B	B	B	B	B	A	B	B	75	Yes
8	A	B	B	B	B	B	B	B	A	A	B	B	B	B	A	A	68	Yes
9	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	100	Yes
10	B	B	B	A	B	B	B	B	B	B	B	A	A	A	B	A	68	Yes
11	A	B	A	A	B	B	B	A	A	B	A	B	A	B	A	A	38	No
12	B	B	B	B	B	A	A	A	A	B	A	B	A	B	A	A	50	No
13	B	B	B	B	B	B	B	B	B	A	B	A	B	A	B	B	81	Yes
14	B	A	B	B	B	B	A	B	B	B	B	B	B	B	B	B	87	Yes
15	A	B	A	B	B	B	A	A	A	B	A	B	B	A	A	B	50	No
16	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	A	87	Yes

Test No.

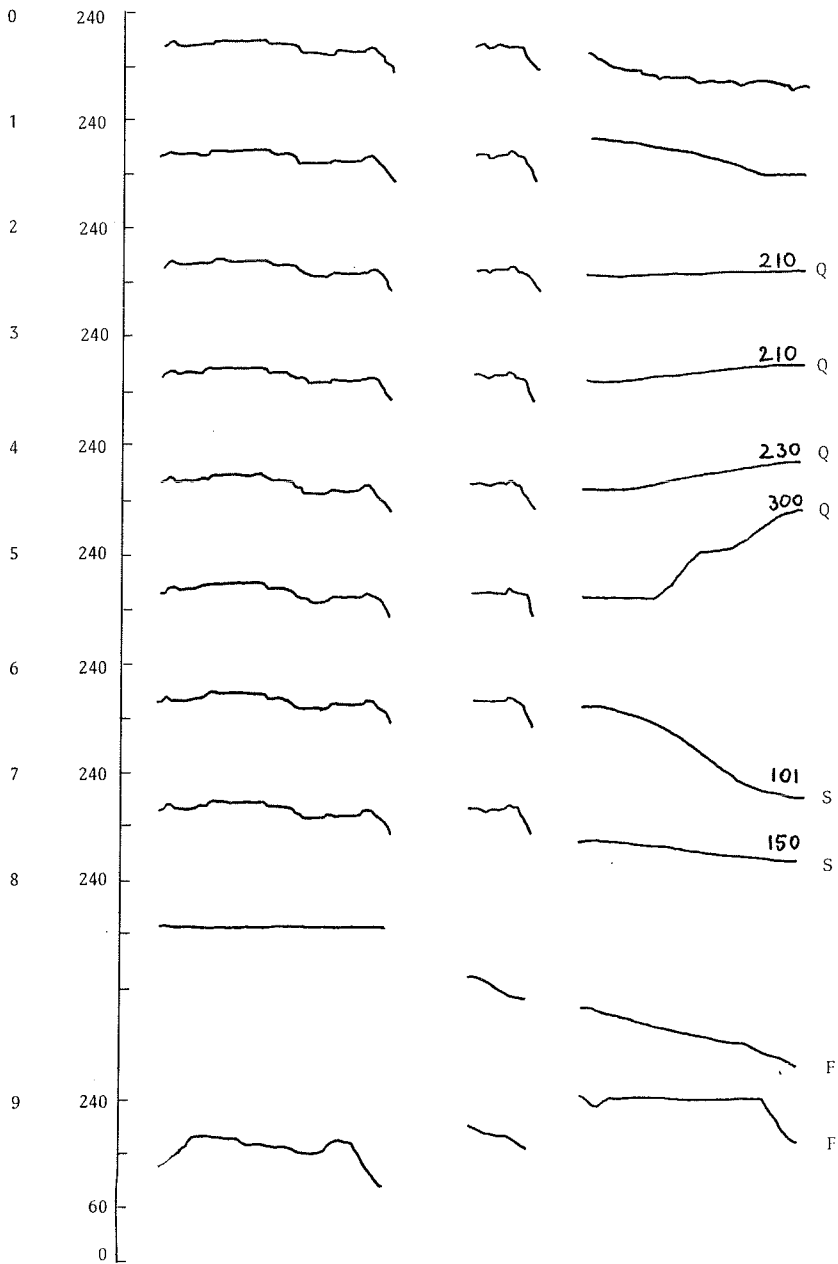


Figure 1. Test sentence: Wāng Yī chōu xiāng yān.

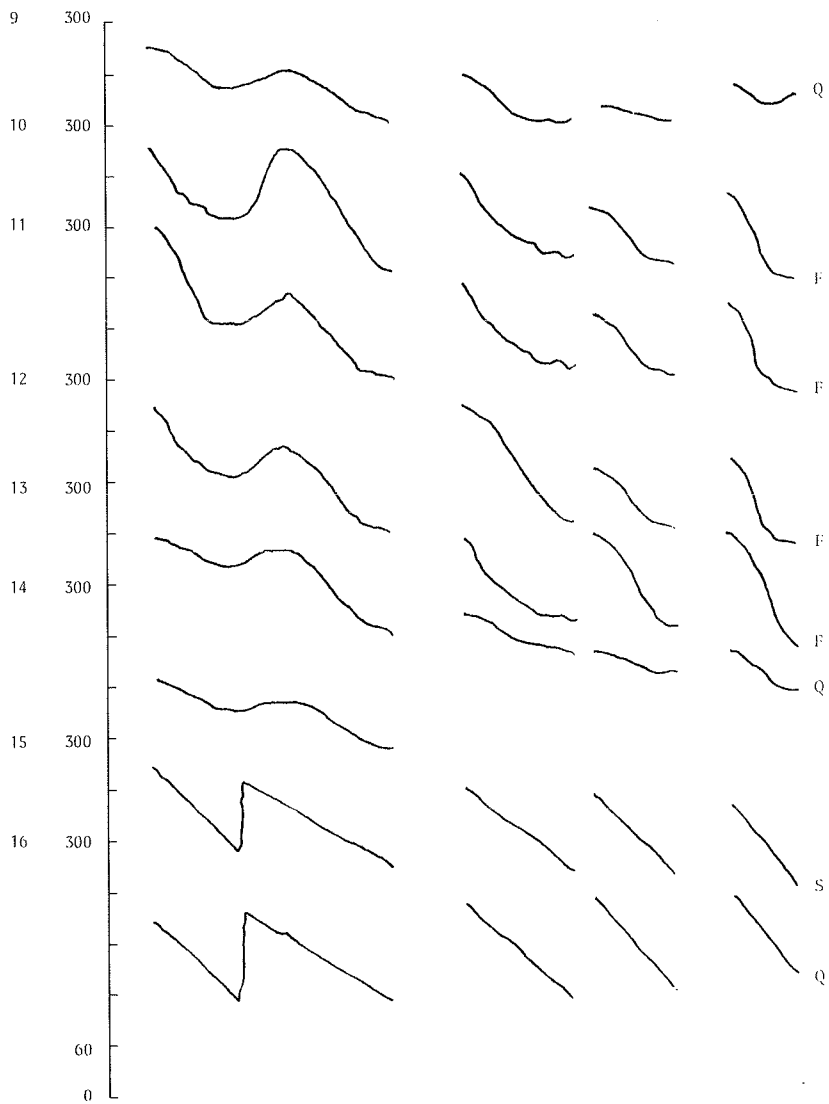


Figure 2 continued

