

Effect of topic structure and sentence length on pause in Mandarin Chinese: Comparing female with male speakers

Qian Wu Bei Wang^{1*} Xiaxia Zhang

¹Minzu University of China

mdwuqian@sina.com, bjwangbei@gmail.com, xiaxia523@163.com

Abstract

This paper studied effects of topic structure and sentence length on acoustic parameters at intonational phrases boundaries, comparing female and male speakers. Twenty native speakers of Mandarin Chinese read 12 short discourses, which contained two sentences. The second sentence was either short or long. And, the transition between the two sentences was either topic continuation, topic elaboration or topic shift, by varying the second sentence. Extensive acoustic analysis showed that [1] for the female speakers, both topic structure and sentence length had significant effects on pause duration and pitch reset, but not on final lengthening, whereas for the male speakers, only sentence length had such an effect. [2] A negative correlation between pause duration and pre-boundary lengthening was found for both the female and the male speakers. However, a positive correlation between pause duration and pitch reset was found only for the females, but not for the males. In general, the female speakers used combined acoustic cues for marking topic structure, whereas the male speakers did not mark topic structure systematically. Effect of sentence length on pause duration is robust while all the other conditions are controlled.

Index Terms: topic structure, sentence length, pause, pitch reset, final lengthening

1. Introduction

The occurrence and strength of prosodic boundaries are influenced by several factors, such as discourse structure [1], syntactic structure [2, 3], phrase length [2-5], and speech rate [6], etc.. Individual difference on pause duration has also been reported in [7] and [8].

It has been found that longer phrases preceding and following a boundary lead to longer pause duration [2-5]. Zvonik and Cummins [5] reported that pause less than 300 ms mostly occurred when the preceding and following phrases were less than 10 syllables, whereas when the pause was longer than 300ms, no clear linear relationship was found between pause duration and phrase length. In addition, it has been found that phrases at pre-boundary and post-boundary have independent influence on pause duration [5]. Ferreira [3] showed that the preparing time before uttering a sentence increased with the number of phonological words.

Topic structure also affects prosodic boundaries [1]. One approach in analyzing topic structure is to make a distinction between sentence topic and discourse topic [9, 10]. Sentence topic represents the proposition as to what information is given

at a local level, whereas discourse topic represents what a whole text is about. It has been found that higher pitch and greater pitch range occur at a higher level in the discourse hierarchy [10]. And lower F0 peak has been found to be in sentences at the end of a discourse [9]. Another approach focuses on topic transitions between sentences, which are categorized into four classes by Nakajima and Allen [11], that is, topic shift, topic continuation, topic elaboration and speech-act continuation. Smith has found that topic shift increases final lengthening and pause duration [1].

Individual difference has been found on marking prosodic boundaries. Whiteside [7] showed that women tend to pause more often than men. And Samuelsson [8] reported that men preferred to use F0 shifts, while women preferred to use pauses and pre-boundary lengthening to mark boundaries.

From above, we can see that acoustic cues at prosodic boundaries are affected by both sentence length and topic structure. However, to our knowledge, there has not been any experimental study investigating the interaction between these two factors, which serves as the goal of the current study. Besides, gender difference will also be systematically studied.

2. Method

2.1 Reading materials

Twelve short discourses were constructed, with two sentences in each discourse. Two factors were manipulated: length of post-boundary sentence (short vs long) and topic structure (continuation, elaboration or shift). Under the framework of Nakajima and Allen [11] topic structure was manipulated according to the semantic transition between the two sentences. In the condition of *topic continuation*, the second sentence talked about the same entity discussed in the previous utterance. The subject of the second sentence was omitted because it is the same as the first sentence. In the condition of *topic elaboration*, the second sentence added some relevant information to the previous statement, and had a different subject, but still talked about the same issue as the first one and the discourse topic was not changed. And for the condition of *topic shift*, not only the subject of the second sentence was changed, but also the discourse topic was changed. The two sentences talked about two totally different events. An example is presented in (1a) to (3a).

(1a) Topic continuation

Qiang leng kong qi xi ji nan fang, duo ci po shi xue xiao ting ke. (Strong cold air appeared in South-China, repeatedly and enforced the school to suspend classes.)

(2a) Topic elaboration

Qiang leng kong qi xi ji nan fang, gong lu duo chu yan zhong jie bing. (Strong cold air appeared in South-China repeatedly and the highway froze seriously.)

(3a) Topic shift

Qiang leng kong qi xi ji nan fang, jing hu gao su zheng shi kai tong. (Strong cold

* Corresponding author.

air appeared in South-China repeatedly. Peking-Hu highway was used for the first time.)

Sentence length was manipulated by adding modifiers before the object, while keeping all the other words the same in the second sentences (see 1b to 1c).

(1b) Topic continuation

Qiang leng kong qi xi ji nan fang, duo ci po shi shi nei suo you xue xiao ting ke.
(Strong cold air appeared in South-China repeatedly and enforced many schools to suspend classes.)

(2b) Topic elaboration

Qiang leng kong qi xi ji nan fang, gong lu duo chu lian xu shu ri yan zhong jie bing.
(Strong cold air appeared in South-China repeatedly and the highway froze seriously for a long time.)

(3b) Topic shift

Qiang leng kong qi xi ji nan fang, jing hu gao tie jin ri quan xian zheng shi kai tong.
(Strong cold air appeared in South-China repeatedly. Peking-Hu highway was used recently for the first time.)

Two base discourses were constructed with consideration of lexical tones before and after the boundaries. We used words with HH+HH and HL+LL. In total, there were 2 (length) × 3 (topic structure) × 2 (tonal condition) = 12 discourses.

To measure final lengthening of the last word in the first sentence, a similar method as in [1] was used. In order to compare the word with a non-sentence-final word, a control sentence was constructed with the target word at a sentence-medial position. Besides, 18 short discourses with different lengths and syntactic structures were used as fillers. All together, there were 12 (experimental discourses)+8(control discourses)+18 (filler discourses)=38 discourses, which made one block. All the discourses were repeated 5 times with random order for each block and each speaker.

2.2. Participants

Twenty speakers took part in the experiment, 10 male and 10 female. They were college students of Minzu University of China. They were native speakers of Mandarin Chinese, and were paid a small amount of money for their participation. All the speakers came from north China.

2.3. Recording procedure

All the speakers were recorded in the speech lab at Minzu University of China, using Cool Edit 2.0 software and a Rode NT1-A microphone. The recording was saved at 22.5 kHz sampling rate to a DELL computer as WAV files.

Before the recording, the speakers were asked to get familiar with all the reading materials. When the experimenter (the first author) noticed any errors, she would ask the speaker to read the discourse again. The recording lasted about 45 minutes.

2.4. Acoustic measurements

The wav files were labeled in Praat with the script of Xu [12]. The measurements included pause duration, the amount of final lengthening in the last word of the first sentence and pitch reset of maximum and minimum F0 at the boundaries. Pitch reset was calculated as the F0 differences between the two words before and after the boundaries. The maximum F0s of the boundary words were calculated, then the difference between them was calculated as the pitch reset of maximum F0.

Similarly, the pitch reset of minimum F0 was calculated as difference of minimum F0 of the boundary words. The durational lengthening of the final word of the first sentence was calculated as the durational difference of the word in the first sentence and the same word in the controlled sentences.

3. Results

3.1. Pause duration

Fig.1 presents the averaged pause duration in different conditions of the male and the female speakers.

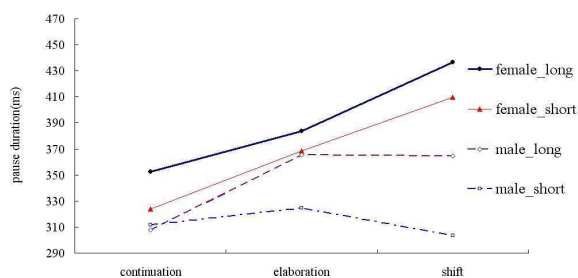


Figure 1. Pause duration in all the conditions

The first observation in Fig.1 is that pause duration of the females (379ms SD=147.5) is longer than that of the males (329ms, SD=80.8). The result of a one-way ANOVA shows that the difference is significant ($F(1,118)=5.208, p<.05$).

With a closer look, we can see that for the females, pause duration is longer when the second sentence is long. As for the topic structure, pause duration is the longest when topic shifts, and pause is the shortest in the condition of topic continuation, with topic elaboration in between. For the males, in the condition of topic continuation, pause duration does not show clear difference between the long and short sentences. But in the condition of topic elaboration and topic shift, pause is longer when the second sentence is long. In addition, pause duration between topic elaboration and topic shift does not differ much for the males.

To test the above observations, a two-way repeated measure ANOVA was carried out with topic structure and sentence length as the independent variables. Since the female and male speakers showed different patterns of pause duration, the statistic tests were carried out separately for them (see Table 1).

Table 1. The *F* and *p* values of the two-way repeated measures ANOVA on pause duration

Gender	Topic F(2, 18)	Length F(1, 9)	Interaction F(2, 18)
Female	4.19*	9.08*	0.04, <i>n.s.</i>
Male	3.12, <i>n.s.</i>	26.25**	4.50*

Note: * stands for $p<.05$, ** for $p<.01$ and *** for $p<.001$.

The results in Table 1 showed that sentence length has an effect on pause duration for both the male and the female speakers. However, the effect of topic structure exists only for the female speakers.

One concern here is that the difference on pause manipulation between the male and the female speaker might be because the speaking rate is different between them. In order to exclude the effect of speaking rate on pause duration, the percentage of pause duration in the discourse duration was

examined (see Fig. 2).

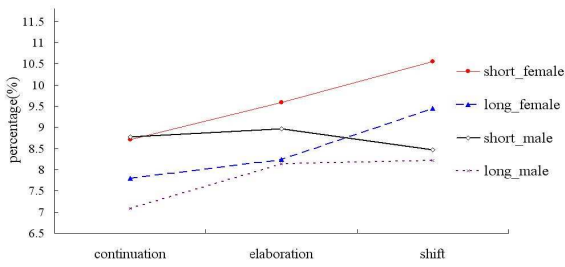


Figure 2. Percentage of pause duration in the discourse

We can see in Fig. 2 that pause percentage is still higher for the females than for the males. The results of a two-way repeated measures ANOVA confirms that both sentence length and topic structure have effects on the percentage of pause duration for the females. And there is no interaction between them. However, only sentence length has an effect for the males (see Table 2). These results are in accordance with those for pause duration.

Table 2. The *F* and *p* values of the two-way repeated measures ANOVA on percentage of pause duration

Gender	Topic F(2, 18)	Length F(1, 9)	Interaction F(2, 18)
Female	3.78*	32.51***	0.70, <i>n.s.</i>
Male	1.68, <i>n.s.</i>	32.82***	4.05, <i>n.s.</i>

3.2. Final lengthening

Fig. 3 shows durational lengthening of the last syllable of the first sentence in all the conditions for the male and female speakers. As can be seen in Fig. 3, the female speakers show greater lengthening than the male speakers. However, the difference does not reach statistical significance ($F(1, 18)=2.09, n.s.$). The results of a two-way repeated measures ANOVA, with topic structure and phrase length as the independent variables, show that none of the variables has a main effect on final lengthening for both female and male speakers.

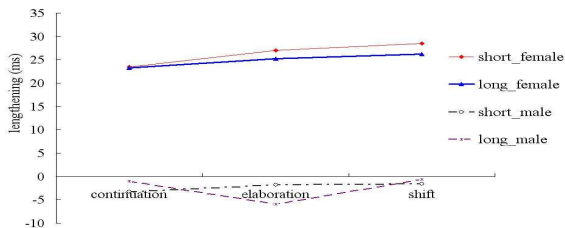


Figure 3. Duration lengthening of the last syllable in the first sentence in different conditions

3.3. Pitch reset

First of all, we need to mention that the second sentence in the condition of topic continuation starts with an adverb, which is usually emphasized, whereas in the other two conditions, the second sentence starts with a noun which is usually not emphasized. For this reason, the condition of topic continuation is not comparable to the other two topic conditions in terms of F0. Here, we will only analyze the conditions of topic elaboration and topic shift.

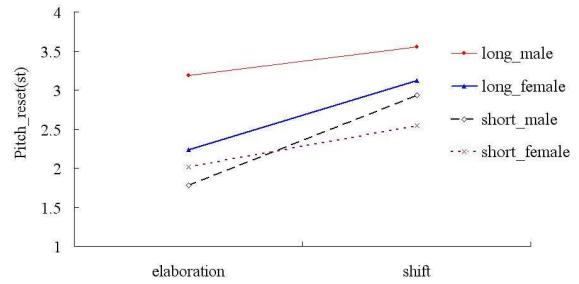


Figure 4. Pitch reset of maximum F0 at the boundary

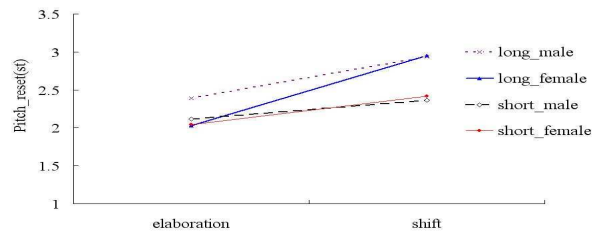


Figure 5. Pitch reset of minimum F0 at the boundary

Fig.4 and 5 present averaged pitch resets of maximum and minimum F0 in four conditions for the male and female speakers. Results of statistic tests are present in Table 3. For the females, both sentence length and topic structure have a main effect on the pitch reset of maximum and minimum F0. The pitch reset is significantly larger when the post-boundary sentence is long and when topic shifts. However, for the males, only significant effect of sentence length is found in maximum and minimum F0. In addition, for both the females and the males, sentence length and topic structure have no interaction.

Table 3. The *F* and *p* values of the two-way repeated measures ANOVA on maximum and minimum F0

	Gender	Topic F(2, 18)	Length F(1, 9)	Interaction F(2, 18)
MaxF0	Female	7.35*	13.80**	1.45, <i>n.s.</i>
	Male	4.55, <i>n.s.</i>	19.63**	4.08, <i>n.s.</i>
Min F0	Female	7.63*	7.17*	1.77, <i>n.s.</i>
	Male	3.49, <i>n.s.</i>	11.32**	1.76, <i>n.s.</i>

3.4. Correlations among the acoustic cues

Table 4 shows correlations between pause duration and final lengthening in all the conditions. We can see that in the condition of topic continuation and elaboration, these two variables have a low negative correlation (about -0.3), whereas in the condition of topic shift, there is no significant correlation. These hold for both male and female speakers.

Table 4. The correlation between pause duration and final lengthening in all conditions

Conditions	Female	Male
Con_S	$r=-0.351***$	$r=-0.222*$
Ela_S	$r=-0.369***$	$r=-0.277**$
Shif_S	$r=0.069, n.s.$	$r=-0.122, n.s.$
Con_L	$r=-0.288**$	$r=-0.309**$
Ela_L	$r=-0.275**$	$r=-0.09, n.s.$
Shif_L	$r=0.021, n.s.$	$r=0.027, n.s.$

Note: Con, Ela and Shif stand for Topic continuation, elaboration, and shift respectively. S and L stand for short and long phrase respectively.

Table 5 shows correlations between pause duration and pitch reset of maximum and minimum F0. For the females,

these two variables have positive correlations. However, for the males, no systematic correlation was found.

Table 5. *The correlation between pause duration and maximum and minimum F0 reset*

	maximum F0(r)		minimum F0(r)	
	Female	Male	Female	Male
Con_S	0.319***	0.258**	0.462**	0.105, <i>n.s.</i>
Ela_S	0.351***	0.094, <i>n.s.</i>	0.263**	-0.028, <i>n.s.</i>
Shif_S	0.357***	0.000, <i>n.s.</i>	0.254*	-0.047, <i>n.s.</i>
Con_L	0.387***	0.281**	0.219*	0.156, <i>n.s.</i>
Ela_L	0.327***	0.173, <i>n.s.</i>	0.149, <i>n.s.</i>	-0.013, <i>n.s.</i>
Shif_L	0.304**	-0.005, <i>n.s.</i>	0.015, <i>n.s.</i>	0.025, <i>n.s.</i>

In summary, final lengthening and pause duration have a low negative correlation in the conditions of topic continuation and topic shift for both the male and female speakers. And a positive correlation between pause duration and pitch reset exists for the females but not the males.

4. Discussion

The current experiment found that both sentence length and topic structure have effects on pause duration and pitch reset, but not on pre-boundary lengthening. In addition, the effects of these two factors show gender differences in that female speakers mark topic structure with F0 variation and silent pause, whereas male speakers do not mark topic structure systematically. For both male and female speakers, the effect of sentence length on silent pause applies.

As found here, longer phrase length after a boundary leads to longer pause duration. These results are in consistent with the previous studies[2-5]. The effect holds for both female and males speakers.

As for topic effect on pause duration, gender difference was also found as the females show such an effect, but the males do not. Moreover, for the male speakers, there is an interaction between topic structure and phrase length on pause duration. The effect of phrase length applies in the conditions of topic elaboration and the topic shift, but not topic continuation. No interaction between topic structure and phrase length was found for the female speakers, though. The above pattern is also found in the percentage of pause duration in the discourse duration. It indicates that the above findings about pause duration is probably irrelevant to speaking rate.

Interestingly, final lengthening does not show any difference in varied topic conditions and phrase lengths for both the male and the female speakers.

The effect of topic structure and sentence length on pitch reset is similar to that of pause duration. For the female speakers, both factors have effects on pitch reset, whereas only sentence length shows an effect for the male speakers. No interaction between the two factors was found for both the female and the male speakers.

5. Conclusions

Putting together these results, we argue that the effect of

sentence length on acoustic variation at an intonational phrase boundary is mostly due to articulatory constrains, so that it is found in both male and female speakers. In contrast, topic structure as reflected in prosody is probably expressive, and is better marked by female speakers.

In general, female speakers seem to use pitch and duration together to mark prosodic boundary, whereas no clear correlation between these two parameters was found for male speakers.

6. Acknowledgement

We are thankful to Tianlai Data Center for the preliminary speech data labeling. We want to thank Qifan Ding for her nice work on speech data annotation. We are grateful to Yi Xu and Fred Cummins for their constructive suggestions.

This study was funded by Nature Science Foundation of China (NSFC) (Nr. 60905062) and “111” Project of Minzu University to the second author.

7. References

- [1] C. L. Smith, "Topic transitions and durational prosody in reading aloud: Production and modeling," *Speech Communication*, vol. 42, pp. 247–270, 2004.
- [2] F. Grosjean, "The patterns of silence: Performance structures in sentence production," *Cognitive Psychology*, vol. 11, pp. 58-81, 1979.
- [3] F. Ferreira, "Effects of length and syntactic complexity on initiation times for prepared utterances," *Journal of Memory and Language*, vol. 30, pp. 210-233, 1991.
- [4] E. Strangert, "Pausing in texts read aloud," in *the XIIIth International Congress of Phonetic Sciences*, Université de Provence, 1991.
- [5] E. C. Zvonik, F. , "The effect of surrounding phrase lengths on pause duration," in *Eurospeech*, Geneva; Geneva, Switzerland, 2003.
- [6] J. Fletcher "Some micro and macro effects of tempo change on timing in French," *Linguistics* vol. 25, pp. 951–967, 1987.
- [7] P. Whiteside, "Temporal-based speaker sex differences in read speech: Some evidence for speaker sex differences," *Journal of the International Phonetic Association*, vol. 26, pp. 23-40, 1996.
- [8] Y. Samuelsson, "Gender effects on phonetic variation and speaking styles." A literature study, 2006.
- [9] N. Umeda, "'F0 declination" is situation dependent," *Journal of Phonetics*, vol. 10, pp. 279–290, 1982.
- [10] I. Lehiste, "The phonetic structure of paragraphs," in *Structure and Process in Speech Perception*, N. Cohen A, S. , Ed., ed Berlin, 1975.
- [11] S. Nakajima, Allen, J., "A study on prosody and discoursestructure in cooperative dialogues," *Phonetica*, vol. 50, pp. 197–210, 1993.
- [12] Y. Xu, "Prosody Pro.praat. Availablefrom," in <http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/S>, ed, 2005-2010.