BREAK WITH FILLED PAUSE IN STANDARD CHINESE AND F0 RANGE OF TONE--A F0 NORMALIZATION IN UTTERANCE

Lin
Institute of Linguistics, Chinese Academy of Social Sciences, Beijing

ABSTRACT

In Standard Chinese, There are two types of break with filled pause and break with silent pause. In this paper, the break with filled pause is mainly discussed.

The major break and the minor break with filled pause, just as that with silent pause are found.

The break with filled pause are most obviously characterized with the lengthening syllable preceding the break and the F0 reset of the tone following the break relative to that preceding the break. But, to differentiate the major break from the minor break with filled pause is not by the lengthening syllable preceding the break and F0 reset of the tone following the break relative to that preceding the break.

Following the F0 normalization formula proposed in this paper, the J values of F0 in each tone of utterances are calculated. It is further found that the major break with filled pause and the minor break with filled pause could be differentiated by the F0 range of the tone preceding the break. The F0 range of the tone preceding the major break with filled pause is wider than that preceding the minor break with filled pause.

It is well known that in Standard Chinese, the F0 range in the syllable is related to sentence stress. Therefore, the syllable preceding the major break is more stressed than that preceding the minor break.

1.INTRODUCTION

From previous research, it is known that the speakers use the prosodic cues, pause and pitch, to audibly structure their spoken message. Listeners, on the other hand, use these phonetic cues to determine the degree of disjuncture in the flow of speech, which supposedly help them to process the meaning of the utterance [1].

The prosodic structures of Standard Chinese have been studied more and more deeply. It is found [2] that there is a systematic difference in temporal and frequency parameters of the syllables preceding the word, word group, phrase, clause and sentence. It is found from natural continuous speech [3] that in Standard Chinese, there are also two types of prosodic phrases: major phrase (MAP) and minor phrase (MIP); The final syllable in MAP is longer than that in MIP without considering tone; The situation is much more complicated when tone is considered. It is further found [4] that the speakers use two strategies to cause the different kinds of breaks: break with silent pause and break with filled pause.

The research reported here focused on what cues are used to differentiate the major break with filled pause from minor break with filled pause. In the following, the methodology used was first introduced, followed by results, conclusion and discussion.

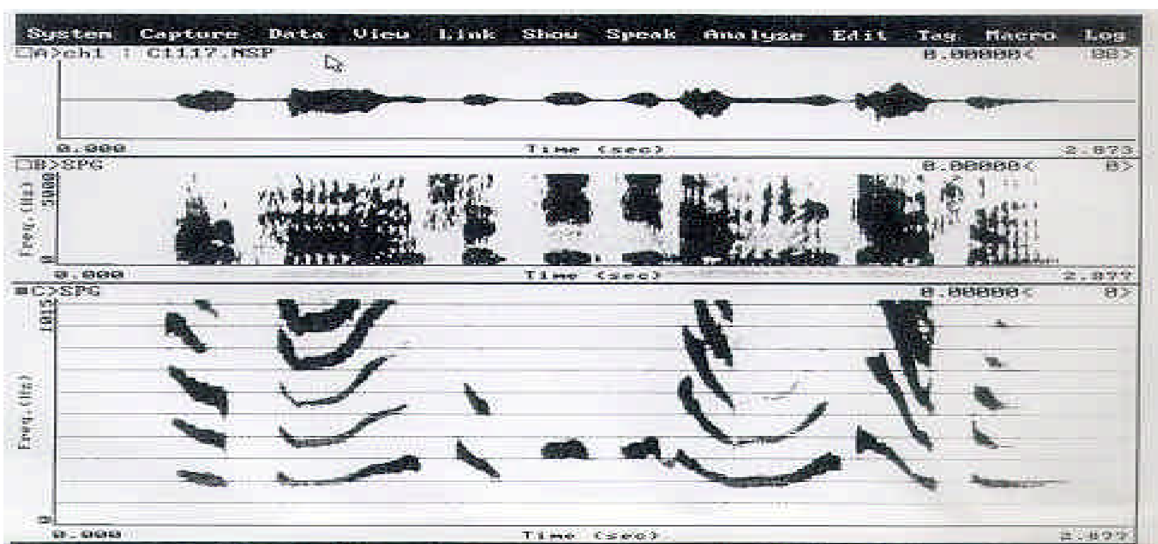

Figure 1 The spectrograms of / ər˩ paŋ˥ ʂʅ˥˩ tɕi˥ tɕi˥ tə˧ miŋ˥ tɕiao˥˩ ʂəŋ˥ / (耳旁是唧唧的鸣叫声) uttered by speaker W

X lengthening Y shortening			X and Y lengthening		X shorting Y lengthing	X as neutral tone	X and Y Shorting
F_0 reset of Y relative to X		no F_0 reset ofY relative to X	F_0 reset of Y relative to X	no F_0 reset of Y relative to X	F_0 reset of Y relative to X	F0 reset of Y relative to X	F_0 reset of Y relative to X
"//"	"/"	"/"	"/"	"/"	"/"	"/"	"/"
X>(T+2σ) Y<(T) δ F_0=138Hz 100Hz 71Hz 66Hz X>(T+σ) Y<(T) x>(T) Y<(T) δ F_0=54Hz	X>(T+2σ) Y<(T-σ) δ F_0=104Hz 39Hz x>(T+2σ) Y<(T) δ F_0=112Hz 86Hz 78Hz 20Hz X>(T+σ) X<(T) δ F_0=139Hz 138Hz 132Hz 109Hz 100Hz 77Hz 70Hz 66Hz 49Hz 43Hz 33Hz 31Hz 15Hz X>(T) Y<(T) δ F_0=116Hz 94Hz 6Hz 74Hz 35Hz 18Hz	X>(T+σ) T<(T) x>(T+σ) Y<(T) x>(T+σ) Y<(T)	X>(T+2σ) Y>(T+σ) δ F_0=30Hz X>(T+2σ) Y>(T) δ F_0=-11Hz X>(T+σ) Y>(T+2σ) δ F_0=86Hz X>(T+σ) Y>(T+σ) δ F_0=50Hz X>(T+σ) Y>(T+σ) δ F_0=41Hz 19Hz 4Hz -15Hz X>(T) Y>(T+2σ) δ F_0=100Hz X>(T) Y>(T) δ F_0=57Hz X>(T) Y>(T) δ F_0=23Hz	X>(T+σ) Y>(T)	x<(T) Y>(T+σ) δ F_0=63Hz X<(T) Y>(T) δ F_0=125Hz 20Hz	x<(T) Y>(T) δ F_0=105Hz x<(T) Y<(T) δ F_0=11Hz 0Hz	x<(T) Y<(T) δ F_0=-32Hz

Table 1 F_0 and duration of the syllables preceding and following the break with filled pause

2 Methodology

44 sentences uttered by a female speaker W of Beijing Mandarin were used in this study. These sentences were taken from the speech database of Standard Chinese [5]. This material had met the following two conditions. The speaker was naive with the purpose of this experiment. And these sentences included a variety of syntactic structures, to anticipate the likelihood that the material would have a variety of prosodic boundary types and degrees of their strength [6]. The shortest sentence had three words, or had two word boundaries, while the longest sentence had nine words, or eight word boundaries.

The data of F0 and duration T in each syllable in utterances were measured from the spectrograms as shown in Figure 1.

Two listening tests were conducted. The aim of the first listening test was to determine where in each utterance was the major break, and where was the minor break. The second listening test was to decide whether the degree of the strength of the major break with filled pause had the same degree oh the strength of the major break with silent pause.

For the first test, on the test tape, each of the 44 utterances was repeated 7 times with 2-second interval to make a triad, each of which was indexed a serial number, and the interval between successive triad was 10 seconds. Each of the listeners was provided with a sentence-list in Chinese characters. For each utterance, the listeners were required to first mark with a slash "/ " between the preceding and the following syllables, if she or he felt a break there; then mark with a double-slash "// " between them, indicating a more salient break there than breaks sensed. Following the Rotondo method [7], The "// " marked by over 70% of the listeners was considered the major break. The "/ " marked by over 50% was considered the minor break. In the first listening test, two types of breaks are found: break with silent pause and break with filled pause.

In the second test, a major break with silent pause judged by all listeners in the first test was considered as a standard major break, and a minor break with silent pause judged by all listeners was considered as a standard minor break. In the second test, these two standard breaks with silent pause were used for comparison with each break of utterances marked in the first test. The listeners firstly listened the two standard breaks with silent pause , then listened each utterance to judge which break was major break, and which was minor break. The two standard breaks with silent pause and the utterance remained to be judged were stored in the computer. The listeners heard to the material over high-quality headphone in an interactive computer session. They could listen to the standard breaks and the utterance as often as they wanted.

In order to get the prosodic boundary with accuracy, the utterances should be delexicalized. Delexicalization can be achieved in various ways. One of the methods was the spectral inversion technique used by Lehiste and Wang [8]. Recently, Pijper and Sandeman [6] reported the result of their experiment that naive and untrained listeners were asked to judge the prosodic boundary in normal speech could make that the judgement of prosodic boundary were not biased by lexical, syntactic and semantic information available to them. In the first test, 27 naive and untrained female listeners were asked to

	tone-1		tone-2		tone-3		tone-4	
break	"//"	"/"	"//"	"/"	"//"	"/"	"//"	"/"
mean J	10.90	7.55	2.10-6.63	3.34-5.02	0.3-0-0.00-1.00	0.5-0.3-2.04	12.00-.00	8.62-2.14
t test of J	t(5)=2.82, p<0.05		t(16)=5.14, p<0.001		t(6)=1.62, p<0.005		t(10)=1.92, p<0.05	
mean duration	215ms	219ms	305ms	249ms	241ms	200ms	215ms	193ms
t test of duration	t(5)=0.83, p>0.5		t(16)=1.69, p=0.1		t(6)=0.17, p>0.5		t(10)=2.07, p>0.5	

Table 2 Average J values of F_0 and average duration of the lengthening syllables preceding the breaks

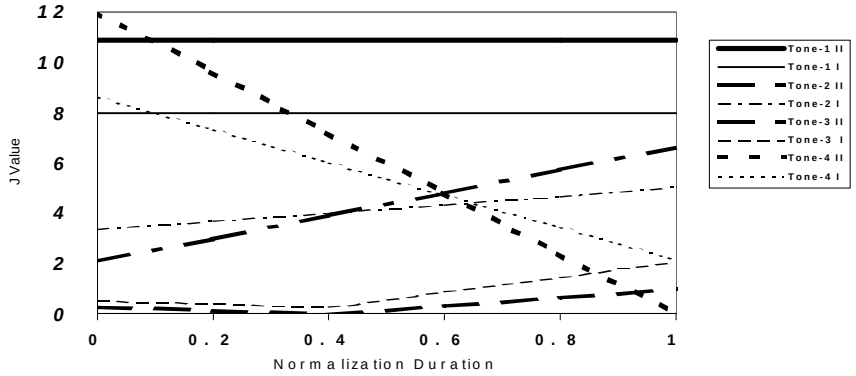

Fig 2. F0 range in tones preceding the major break and minor break

judged the breaks in the utterances, of them 17 listeners studying in City University of Hong Kong and 10 working in Beijing..Only the 10 listeners working in Beijing participated in the second test.

In order to compare the F0 range in each syllables related of each utterances, we proposed the following formula to calculate the J values of F0:

J=12*(log F0 / F0 min) / (logF0 max/ F0 min)

Here, F0 max denotes the maximum value of F0 in any utterance
the any F0 of the utterance. J is the normalization value of F0: The maximum value is 12, but the minimum value is 0.

3. RESULTS

Table 1 gives F0 and duration of the syllables preceding and following the breaks with filled pause. In the table, X and Y denote the syllables preceding and following the break respectively. T refers the mean duration of the syllables in each utterance, σ its deviation. (T±σ), (T±2σ),(T±3σ) is used to represent the degree of lengthening and shortening of the syllables; δF0 denotes the F0 reset of Y relative to X, and δF0 means a F0 change between the two syllables that goes beyond the normal F0 range for the tones there.

Table 1 shows that in 44 utterances, there were 54 breaks with filled pause. There were 5 types of temporal patterns of X and Y: X lengthening and Y shortening, X and Y lengthening, x shortening and Y lengthening, X as neutral tone, and X and Y shortening. However, the lengthening syllables preceding the breaks were 87% of all.

In the 54 breaks with filled pause, there were two types of F0 reset: F0 reset and no F0 reset, but 93% of the 54 breaks had F0 reset between the syllables preceding and following the breaks.

Based on the above analysis, it was shown that the break with filled pause was most obviously characterized with the lengthening syllable preceding X and the F0 reset of Y relative to Y.

In the 54 breaks with filled pause, there were 7 major breaks, and all of them were that the syllables preceding the breaks were lengthened. In second test, the most listeners judged that the degree of the strength of the 7 major breaks with filled pause had the same degree of the strength of the major break with silent pause.

Table 1 also shows that in the conditions of the lengthening X and the F0 reset of Y relative X, there were not only 7 major break with filled pause, but also 25 minor break with filled pause. and in the condition of X<(T+2σ) but Y<(T), and X relative to Y has F0 reset, both of major break and minor break were caused. These evidences lead to that to differentiate the major break from the minor break with filled pause is not by the lengthening syllable preceding the break and F0 reset of the tone following the break relative to that preceding the break.

What cues can differentiate the major break from the minor break with filled pause? Following the F0 normalization formula above, the J values of F0 in the lengthening syllables preceding the major and minor breaks were calculated.. Table 2 gives the mean J value of F0 and mean duration T of the lengthening syllables preceding the breaks and their result of the t test. In table 2, the thick lines refer F0 range in tones preceding the major break, the think line the minor break. Table 2 shown that their duration can not differentiate between the major break and minor break, because the duration is no significant at p= 0.05. But At p=0.05, the F0 range in tones can differentiate the major break from the minor break. Therefore, we think that the F0 range preceding the breaks can differentiate between the major break and minor break with filled pause. F0 range in tones preceding the major break is larger than that preceding the minor break. For example, Fig. 1 shown the spectrograms of the utterance “er pang shi ji ji te ming jiao sheng” uttered by thr speaker W (pingyin symbol in Standard Chinese). A major break occurred between the major prosodic phrases “er pang” and “ji ji te ming jiao sheng”, because the F0 range of the syllable “pang” was wider, but a minor break occurred between the minor prosodic phrases “shi ji ji te ” and “ming jiao sheng”, because the F0 range of the syllable “te’ was narrower. Figure 2 shows the F0 range in tones preceding the major break and the minor break.

4. CONCLUSION AND DISCUSSION

In Standard Chinese, There are two types of break with filled pause and break with silent pause. In this paper, the break with filled pause is mainly discussed.

The major break and the minor break with filled pause, just as that with silent pause are found. there are 5 types of the temporal pattern of the syllables preceding and following the breaks with filled pause., but 87% of the breaks with filled pause are that the syllable preceding the breaks with filled pause are lengthening. There are 2 types of F0 reset of the syllable following the break relative to that preceding the break: F0 reset and no F0 reset, but 93% of the breaks with filled pause are that the F0 reset of the tone following relative to that preceding the break occur. Based on the above analysis, it shows that the break with filled pause most obviously characterized with the lengthening syllable preceding the break and the F0 reset of the tone following the break relative to that preceding the break.

In the lengthening syllable preceding the break with filled pause and reset of the syllable following the break relative to that preceding the break, there are not only 7 major breaks, but also 25 minor break with filled pause. Even in the condition that the duration T of the syllable preceding the break is longer than ($T+2\sigma$), and the F0 reset of the syllable following the break relative to that preceding the break occur, both the major break and the minor break with filled pause are caused. These evidences lead to that to differentiate the major break from the minor break with filled pause is not by the lengthening syllable preceding the break and F0 reset of the tone following the break relative to that preceding the break.

Following the F0 normalization formula given above, the J values of F0 in each tone of utterances are calculated. It is further found that the major break with filled pause and the minor break with filled pause can be differentiated by the F0 range of the tone preceding the break. The F0 range of the tone preceding the major break with filled pause is wider than that preceding the minor break with filled pause.

It is well known that in Standard Chinese, the F0 range in the syllable is related to sentence stress. There, the syllable preceding the major break is more stressed than that preceding the minor break.

REFERENCES

[1] Sanderman, A. A. And R. Collier 1996. Prosodic rules for the implementation of phrase boundaries in synthetic speech. *Journal of the Acoustical Society of America*, Vol. 100,pp.3390-3397.

[2] Yang, Yufang 1997. Prosodic cues to syntactic boundarie, *Chinese Journal of Acoustics*, Vol. 22, pp. 414-422.

[3] Li, Aijun 1998. "Durational characterics of the prosodic phrase in Standard Chinese", *The Proc. of the Conference on Phonetics of the Languages in China*, pp. 65-68.

[4] Lin, Maocan 1998, The acoustic manfestation prosodic phrase boundaries in Standard Chinese, *The Proc.of the Conference on Phonetics of the Languages in China*, pp. 69-72.

[5] Zu, Yiqing 1999. The text design for continuous speech database of Standard Chinese, *Chinese Journal of Acoustics*, Vol. 18, No. 1.3

[6] Pijper, Jan Roelof and Angelien A. Sanderman 1994. On the perceptual strength of prosodic boundaries and its relation to suprasegmental, *Journal of the Acoustical Society of America*, Vol. 96, No. 4.

[7] Lihiste, L. And W. S.-Y. Wang 1977. Perception of sentence and paragraph boundaries with and without semantic information, In *Phonologica* 1977, edited by Dressler and Pfeiffer (Institute fur Sprachwissenschaft der Universitar Innsbruck, I nnsbruck), pp. 277-283.

[8] Swert, Marc 1997. Prosodic features at discorse boundaries of different strength, *Journal of the Acoustical Scociety of America*, Vol. 101, No. 1.