

## PROSODIC STRUCTURE AND HIERARCHICAL STRESS IN UTTERANCE OF STANDARD CHINESE---ONE OF CUES TO CHINESE INTONATION

*LIN Maocan & YAN Jingzhu*

Phonetics Laboratory, Institute of Linguistics, CASS, Beijing

Maocanlin@263.net

### ABSTRACT

Prosodic word and its prominence and prosodic phrase are examined in this experiment. And the prominence in prosodic word is related to stress. It seems to us that the hierarchical stress in sentence spoken is one of intonational cues in Chinese. Tone and intonation in Chinese are two different phonological events in spoken sentence.

### 1. INTRODUCTION

Since 1980s, Intonation in Chinese has been paid more attention to by scholars using experiments. The basic idea of Wu in studying intonation of Chinese (1980,1990) is that polysyllable tone-sandhi is the invariant units of intonation in spoken Standard Chinese. Shen has found (1985, 1994) that the intonation of Chinese is related to the tonal range of pitch. Susan Shen (1990) claims that there are at least three distinct intonational tunes, which she calls Tunes I, Tunes II, and Tunes III. These are used for yes-no question (Tunes I), wh-question (Tunes II), and alternative ('is it A or B?') (Tunes III). Tseng (1990) suggested the following points in her studying of the relation between tone and intonation: (1) The declination effect does exist in Mandarin Chinese, but only for a very small portion of the sentences studied. (2) The declination theory is not only weak in its predictive power, but also proves to be too strong a claim in terms of its universality. (3) The breath-group theory offers a more plausible alternative for characterizing intonation. (4) The interaction between syntax and intonation does occur, thus further

supporting previous research on intonation. (5) The existence of intonation patterns that can not be explained by either theory should not be ignored. (6) Individual variation in producing spontaneous connected speech as well as in well-formed read sentences poses more difficulty for any proposed intonation theory. (7) The declination theory may prove to be more feasible in characterizing the discourse effect rather than a sentential phenomenon.

Chilin Shi (1997) suggested that Mandarin third tone sandhi provides strong evidence for a hierarchically organized prosodic structure. The phenomenon that allows third tone sandhi across a subject-verb boundary cannot be explained by syntactic juncture, but easily accounted for by giving them the same status in prosodic structure: w(ord) in Beattie and f(oot) in Shih. Since the mid-1980s, all scholars for studying tone sandhi propose that it is the prosodic structure that lies behind the complex facts of Mandarin tone sandhi. The domain of tone sandhi is defined as a small prosodic unit by WANG Hongjun (1999, 250-251). Intonational features of tone and relative prominence are distributed in utterances in ways *allowed by* the prosodic structure (Ladd, 1996, p. 10).

The results and opinions above, especially, the suggestions above by Tseng, Shi' idea of tone sandhi (1997), Wang' definition of the domain of tone sandhi and Ladd' theory of intonation inspire us to examine the rule of prosodic structure in intonation and the hierarchical stress in its sentence

### 2. METHODOLOGY

2.1. Speech material: 59 sentences read by

three native speakers (21 by M00, 21 by M01 and

17 by F00) were used. These sentences were taken from the speech database of Standard Chinese (Zu, 1999).

2.2. listeners: 20 native and untrained listeners who were born in Beijing participated this listening tests.

2.3. Listening tests: One listening test is to decide the boundary of prosodic phrase, another one is to decide what are the prosodic word and its prominence.

Prosodic words were sliced from each utterance by “gating equipment” in Kay Multi-speech model 3700. In each prosodic word, one or two syllables that sounded with acute and intense were judged only by three listeners. The syllable(s) perceived with acute and intense refers to prominent part in prosodic word.

2.4. Acoustic analysis: The data of  $F_0$  and duration (T) of each syllable in utterance were measured from the spectrogram made by Kay Mul-speech model 3700.

2.5. Normalization of  $F_0$  and T: The following formula (Lin 1999,2000) was used to normalize the

$F_0$  :

$$J=5 \times (F_0 - F_{0\min}) \div (F_{0\max} - F_{0\min}),$$

$F_{0\max}$  denotes the maximum value of  $F_0$  for a given sentence,  $F_{0\min}$  the minimum value of  $F_0$  for the

same sentence,  $F_0$  the value that will be measured.

T was normalized using the following formula:

$$d = (t - t_{ave.}) / s_d$$

$t_{ave.}$  denotes the average T of all finals for a given speaker,  $s_d$  its standard deviation;  $t$  denotes the T value of the final that will be measured,  $d$  the normalized T value of the final.

### 3. RESULTS:

3.1 What's prosodic word? Prosodic word refers to those syllable-group that are uttered closely

together judged by the listeners. Prosodic words are divided into two kinds: pure prosodic word and compound prosodic word. Pure prosodic words are referred to those syllable-groups that were uttered closely together judged by more than 85% listeners; compound prosodic words are referred to those syllable-groups that were uttered closely together judged by about 60% listeners.

3.1.1  $F_0$  manifestation in prosodic word:

a. Disyllabic and trisyllabic lexical words including proper nouns, and the preposition and adverbial word emphasized, etc. are pure prosodic word. In pure prosodic word, both the  $F_0$  register of the syllable and the  $F_0$  range of the syllable-group following the tone-1, tone-2 and tone-4 go down and are reduced; the  $F_0$  register of them following tone-3 doesn't go down.  $F_0$  reset always occurs between the prosodic words; the duration in one syllable often longer than other syllables. The syllables with higher  $F_0$  in tone-1, tone-2, tone-4, and with lower  $F_0$  in tone-3 sounded with acute and intense refer to the prominence part in pure prosodic word.

b. Compound prosodic word contains at least two lexical words. The acoustic characteristics are that  $F_0$  register and  $F_0$  range of one syllable-group are lower and narrower than other one. If there is any attachment in it, the manifestation of  $F_0$  in attachment should be considered. The larger  $F_0$  range and register often with acute and intense refers to the prominence in compound prosodic word.

3.1.2. Signals to boundary of prosodic word.

Speakers	Number of Prosodic word	Mean T and its variance of final in the first syllable in prosodic word	Mean T and its variance of final in the last syllable in prosodic word	t test
M00	37	$T_b=149ms$ $\sigma_b=47ms$	$T_e=163ms$ $\sigma_e=48ms$	>0.1
M01	40	$T_b=153ms$ $\sigma_b=47ms$	$T_e=155ms$ $\sigma_e=42ms$	>0.5
F00	37	$T_b=163ms$ $\sigma_b=44ms$	$T_e=164ms$ $\sigma_e=50ms$	>0.1

Table 1 Comparism between the final'durations (T) in the first and last syllables of prosodic word

Speakers	Mean F <sub>0</sub> , its variance and number of positive F <sub>0</sub> reset	mean F <sub>0</sub> , its variance and number of negative F <sub>0</sub> reset
M00	$F_0(\mu)=30.38Hz$ $F_0(\sigma)=15.37Hz$ N=39	$F_0(\mu)=-26.27Hz$ $F_0(\sigma)=-11.75Hz$ N=9
M01	$F_0(\mu)=35.43Hz$ $F_0(\sigma)=20.55Hz$ N=46	$T_f(\mu)=-23.42Hz$ $F_0(\sigma)=-17.35Hz$ N=12
F00	$F_0(\mu)=86.36Hz$ $F_0(\sigma)=59.00Hz$ N=45	$F_0(\mu)=-43.33Hz$ $F_0(\sigma)=-8.15Hz$ N=3

Table 2 F<sub>0</sub> reset between prosodic words

Tables 1 and 2 show that there is no systematic difference between the durations of finals in the last and first syllables of prosodic words, namely, the duration of fianl in the last syllable is not longer than that in the first one in prosodic word. But F<sub>0</sub> reset always occurs between prosodic words. It is a signal to cue to the boundary of prosodic word.

It is the characteristic features of prosodic word that the F<sub>0</sub> reset always occurs between the prosodic words and F<sub>0</sub>s in syllables have their own manifestations.

3.2 What's prosodic phrase? Prosodic phrase refers to those prosodic words that are seperated by more clear breaks with silent or without silent (Lin, 20000). Boundaries of the major prosodic phrases were caused with those breaks that were judged by more than 85% listeners; boundaries of the minor prosodic phrases were caused with those breaks that were judged by 65-85 % listeners.

The signal to boundary of prosodic phrase is as

follows: F<sub>0</sub> reset occurs between the syllable preceding pause with and without silent and that following pause with and without silent. The duration of final in the syllable preceding the pause without silent only is lengthened, its duration is about 200ms.

3.2 Prominence and stress: Figure 1 shows the normalized F<sub>0</sub> and T in each syllable of the utterance “国际航空公司飞上海的航班因大雾取消了”. “国际航空公司飞上海的航班” and “因大雾取消了” are major prosodic phrases, as they are seperated with pause with silent. “国际航空公司” and “飞上海的航班” are minor prosodic phrase, as they are seperated by break without silent that is caused by the lengthened duration of “司” in “公司”. Also, The F<sub>0</sub> manifestation in “国际航空公司” makes it a compound prosodic word. “飞上海的” and “大雾取消了” are compound prosodic words, “航班” and “因” are pure prosodic words.

In “国际航空公司”, “国际”is with more acute

and intense than “航空公司”; In “飞上海的”, “上海” is with more acute and intense; “上海” is also with more acute and intense than “航班”. “大雾” is with more acute and intense than “取消”. “因” also with acute and intense. The prominent part in prosodic word is those one or two syllables that are perceived with acute and intense.

The stress in prosodic word refers to the prominent part. The stress in prosodic phrase is those prominent part that are most prominence, namely most acute and intense syllables in words contained. The stress in utterance is those prominent part that most prominence, namely most acute and intense in prosodic phrase contained. So, “上海” and “因大雾” are the stress in each prosodic phrase, because the  $F_0$  range is wider than that in others. The stress in this utterance may be in “上海”。

### 3. CONCLUSION AND DISCUSSION

In Standard Chinese, prosodic word is a “ $F_0$  variation group”, and  $F_0$  reset always occurs between the prosodic words. It doesn't like prosodic word in English in that the duration of the last syllable

is lengthened (Prince, 1991). The prosodic phrase is characterized not only by pause with silent, but also by pause without silent.

It is the perceived acute and intense syllable(s) (prominence) that is induced by more higher  $F_0$  in syllable and/or more wider  $F_0$  range of syllable-group in prosodic word. The stress in prosodic word refers to the acute and intense syllable(s). So, Prosodic word has its own stress. It is the stress in prosodic phrase that is the most acute and intense syllable(s) or most prominent part in its

prosodic words. Also, it is the stress in utterance that is the most acute and intense syllable in its prosodic phrases. In Standard Chinese, Stress has its hierarchical pattern as shown in Figure 2.

Chinese is a tone language; Tone is lexical specified.  $F_0$  in syllables can be varied to different extent, even to lose its identity, due to the effects of tone sandhi and the perturbation by  $F_0$  coarticulation; The variations in  $F_0$  of syllables are the events that are due to the intersyllabic action, of course,  $F_0$  coarticulation across adjacent syllable. However, rise or down of  $F_0$  register and expansion or contraction of  $F_0$  range is caused by utterance, it is the events that are due to utterance level. It seems to us that lexical tones have been distinguished from pitch range modification in Chinese, namely, tone information has been differentiated from stress pattern in utterance. The hierarchical stress may be one of cues to Chinese intonation. We will further discover another cues to Chinese intonation.

#### References:

- [1] Ladd, D. Robert, *Intonational Phonology*, Cambridge: Cambridge University Press, 1996.
- [2] Lin, Maocan & Yan Jingzhu, “Tonal coarticulation in quadrisyllabic words and phrases with normal stress of standard Chinese”, *Chinese Journal of Acoustics*, Vol. 12, No. 3, 1993;
- [3] Lin, Maocan,  $F_0$  Construction in utterance and its function---Normalization  $F_0$  and duration of each syllable, *Proceedings of Morden phonetics*, (chief editor: Lu Shinan), Jincheng Press, 1999;
- [4] Lin, Maocan, “Break and prosodic structure in utterances of Standard Chinese”, *Contemporary Linguistics*, No. 4, 2000 (in Chinese);
- [5] Price, P. J., Ostendorf M, Shattuck-Hufnagel S., & Fong C., The use of prosody in syntactic disambiguation, *Journal of Acoustics of American Society*, Vol. 90, pp. 2956-2970, 1991;
- [6] Shen, Jong, Pitch range and intonation of Chinese, *Working Papers of Experimental Phonetics*, Beijing: Beijing Press, 1985;

- [7] Shen, Jong, Intonational construction and patterns of intonation in Chinese, *FangYan*, No. 4, 1994;
- [8] Susan Shen, Xiao-nan, The Prosody of Mandarin Chinese, Berkeley: University of California Press, 1990;
- [9] Shih, Chilin, Mandarin third tone sandhi and prosodic structure. In *Studies in Chinese phonology*, eds.: Wang Jialing and Norval Smith, pp. 81-123, Berlin: Mouton de Gruyter, 1997.
- [10] Tseng, Chiu-yu, *An Acoustic Phonetic Study On Tones In Mandarin Chinese*, Special Publications No. 94, Institutw of History & Philology Academia Sinica, Naking, Taipei, Taiwan, 1990.
- [11] Wang Hongjun, *Non-linear Phonology in Chinese-Phonological pattern and monosyllable*, Beijing: Beijing University Press, 1999;
- [12] Wu, Zongji, Tones' variation in sentences of Standard Chinese, *Zhongguo Yuwen*, no. 6, 1980;
- [13] Wu, Zongji, Can polysyllable tone-sandhi be the invariant units of intonation in spoken Standard Chinese? Proc. of 1990 Inter. Conf. On Spoken Language Processing, pp. 18-22, 1990, Kobe, Japan.
- [14] Zu, Yiqing, "Design the data-base for speech recognition", Chinese Journal of Acoustics, Vol. 18, No. 3, 1999.

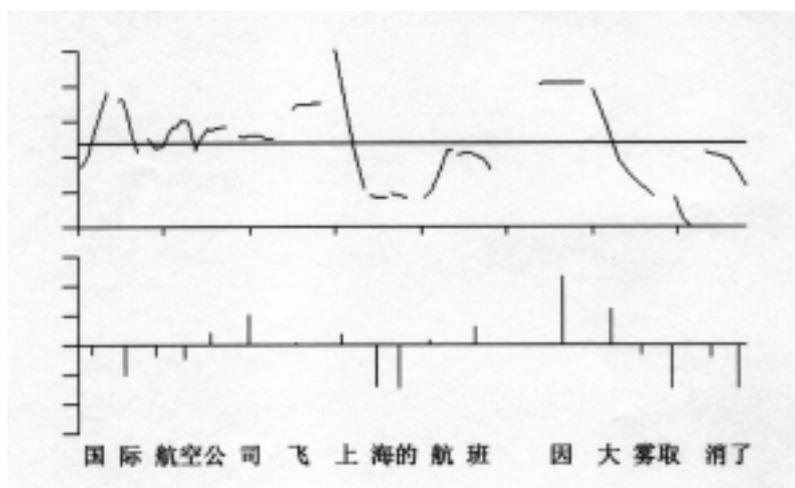


Figure 1 normalized  $F_0$  and T in utterance “国际航空公司飞上海的航班因大雾取消了” uttered by M01

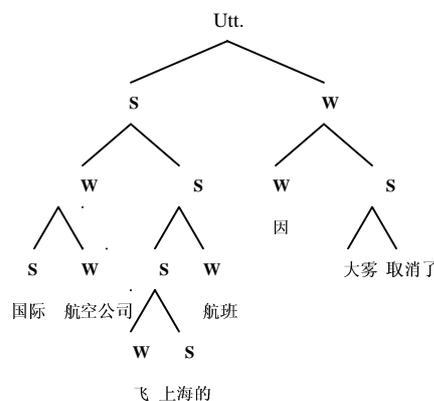


Figure 2 The metrical distribution in utterance “国际航空公司飞上海的航班因大雾取消了” uttered by native speaker M01 of Beijing