Articulatory Strengthening and Prosodic Hierarchy[1]

**Jianfen Cao & **Yuling Zheng*

*Institute of Linguistics & **Institute of Ethnology and Anthropology
Chinese Academy of Social Sciences

***jianfencao@hotmail.com** ****zhyl-cass@263.net**

Abstract

This paper reports a set of results based on the spectral and EPG measurements to the read speech copra in Mandarin Chinese, aim at the observation on the relationship between articulatory strengthening and prosody hierarchy. The data obtained both from acoustic and physiological measurements indicate that, articulatory manifestation of any segment in real speech are closely relevant to their prosodic position or status in connected speech. Therefore, it makes capable to predict the hierarchical organization of speech prosody from the strength of such articulatory strengthening. At the same time, this evidence further reveals the existence of anticipatory planning in speech production. Consequently, our finding should be not only of benefit for Chinese speech processing, but also provides a new angle of view to understand the mechanism of speech production in general.

1. Introduction

Many studies on various languages, such as English, Korean, French and so on, have shown that articulatory strengthening of segments are usually taken place at the edges of prosodic domain, therefore, the articulatory strengthening of segments is a function of the hierarchical level of the prosodic structure [1] [2]. As for Chinese, still lack of approach directly respect in this subject up to date. In the fact, however, some evidences have been observed in several relevant studies. For example, as reported on Chinese rhythm in Cao[3][4], the clear segmental lengthening and pitch prominent does occur at prosodic boundaries and stressed / accented locations. These prosodic location-dependent variables should be regarded as a part of the reflection of articulatory strengthening. To examine whether these findings can be further confirmed by physiological data, an extended approach has conducted by using EPG measurements, in which linguo-palatal contact state for syllable-initial consonants and syllable-final rhymes in different prosodic locations were compared, the obtained data seems well-matched with the findings mentioned above.

The present paper reports our fundamental results, and discusses the prosodic influence on the realization of segment's articulation in Mandarin Chinese.

2. Test Materials and Measurements

Following the popular theory, e. g. [5] and the relevant experimental results obtained from Mandarin Chinese, the prosody structure mainly consists of three layers, i. e., intonation phrase / sentence, prosodic phrase (PP) and

prosodic word (PW), which were determined in terms of perception test. They are organized as a hierarchy and serve as both a grouping and a prominence-marking function in speech [6].

In order to examine the influence of prosodic boundary and stress/accent upon the articulatiory manifestation of segments, two sets of test materials were employed in this study. One set of them consists of 4 speaker's utterance selected from the ASCCD [7], which is a discourse corpus read by multi-speakers. Another set of speech materials were chose from a speech corpus of EPG in Mandarin Chinese [8], which was produced by two speakers who wearing a custom-made false palate embedded with 96 contact electrodes of the Kay Elemetrics EPG system during speaking. Test tokens contain the syllables at the edges of different prosodic domain and the stressed or accented ones in the utterance.

Acoustic measurements, including pitch and duration, were made on the base of first set materials, and the physiological data were measured from the second set of materials. In the latter portion, the measurements contain the data of physiological duration, biggest ratio of lingua-palatal contact area (RCA) and the duration ratio of consonant closure/ or vowel vs. the syllable (DRCS). The target segments to be tested here include syllable-initials (consonants) and finals (rhymes).

3. Results and Discussion

3.1 Acoustic data and analyses

3.1.1 Durational variations

(1) Prosodic boundary-dependent segmental lengthening

Durational measurements based on the first set of materials were briefly summarized in Table 1, in which, the bold figures show a more or less lengthening by comparing to the general mean value in the utterance. At first, if take whole syllable duration as the variable, a post-boundary lengthening appears only at the beginning of sentence level, and a pre-boundary lengthening occurs only at the end of

[1] *Supported by the National Fund of Natural Sciences, NO. 10374117

prosodic phrase (PP) level. However, if we further look at the temporal distribution within the syllables, then we can find that (1) Post-boundary lengthening commonly occurs in all prosodic layers, and the degree of lengthening is cumulative: the higher the level being, the greater the magnitude of lengthening is. In contrarily, pre-boundary lengthening only occurs significantly (p=0.006) in PP level. (2) Post-boundary lengthening is dominant by syllable-initial (consonant), while pre-boundary lengthening is dominant by syllable-final (rhyme, including vowel and the potential nasal ending).

Table 1: *Average duration (ms) of syllables and durational distribution within syllables in different prosodic positions*

location	syllable	Ratio of syll.-initial vs. syllable	Ratio of syll.-final vs. syllable
Post-boundary:			
Sentence level	**216.3**	**0.444**	0.649
PP level	190.2	**0.425**	0.693
PW level	179.0	**0.386**	0.666
Pre-boundary:			
Sentence level	196.6	0.309	0.603
PP level	**235.5**	0.294	**0.962**
PW level	168.1	0.216	0.719
General mean in utterance	*197.6*	*0.319*	*0.733*

(2) Stress-induced segmental lengthening

Generally, there are two main categories of sentence stress in Chinese: grammatical (i.e., default) stress and logical (i.e., marked) stress. According to measured data in this investigation, there seems no clear lengthening phenomenon referred to default stress; while an obvious syllable lengthening can be observed in marked stress case. Moreover, comparing to boundary-induced lengthening described above, such stress-induced lengthening seems relatively balanced over the whole syllable, this situation is similar to that found in English and Japanese by Campbell [9].

Table 2: *Average value(Hz) of pitch level and pitch range in different prosodic positions*

Boundary positions	Pitch level	Pitch range
Sentence level:		
Post-boundary	187.9	32.9
Pre-boundary	113.5	19.1
PP level:		
Post-boundary	180.0	44.2
Pre-boundary	152.0	35.3
PW level:		
Post-boundary	174.4	31.8
Pre-boundary	156.6	26.6
General mean in utterance	166.7	32.4

3.1.2 Pitch variations

(1) Prosodic boundary position-dependent pitch prominent

Table 2 gives the average value of pitch level (i. e., pitch height, which was calculated in terms of the mean of the high point and low point in a syllable's tone pattern) and pitch range in different prosodic positions. As one of the boundary maker, both the pitch level and pitch range of syllables in post-boundary are always higher than that in pre-boundary, no matter in what prosodic levels.

In terms of these variations, a systematic pitch resetting is result at all prosodic boundaries, where the pitch level in post-boundary syllable is generally higher than that in pre-boundary syllable. Note that, however, because of the [+L] register feature of the 3rd tone in Mandarin, when the post-boundary syllable with a 3rd tone, then its pitch level is not higher but lower than that at pre-boundary.

Obviously, it is also a kind of distinctive feature of strengthening in terms of the contrast in pitch prominent, i. e., the H* for domain-initial, and the L* for domain-final, and the L*feature is also referred to the intonational $L^{\%}$ of boundary tone.

(2) Stress-induced pitch accent

As what found in other languages, stressed syllable in Chinese also shows a clear pitch accent. It is generally realized by so-called upstep of pitch level and expansion of pitch range. Table 3 gives an example by summarizing 4-speaker's pitch level referred to a prosodic phrase, where 恢宏 is a PW who loads a phrasal stress, thus, pitch accent occurs in the whole word, and it is true for all speakers' speech as the bold figures shown in this table.

Generally, due to the distinctive features of 4-tone's pitch register in Mandarin, i.e., H for 1st tone, L for 3rd tone , R for 2nd tone and F for 4th tone, the pitch accent in Mandarin follows a rule of "the high one will be the higher, the low the lower", the details can be observed from Fig.1. Specifically, in the case of stressed syllable with the 1st or 4th tone, for example, the /ba1/, /zhen1/ and /tai4/ in the utterance, their pitch are accented as H*; while the pitch accent of a 3rd tone syllable, like the /jia3/ in the utterance, its pitch accent will behave as L*, instead of H* in general, in other words, it is realized through downstep, in stead of upstep of pitch level, and combined with a more complete and typical tone pattern. For example, the original pattern of 3rd tone and 4th tone in Mandarin is LL and HF respectively, therefore, the tone pattern of stressed syllable /jia3/ is obviously prominent than those of unstressed syllables /bai3/ and /mai3/, and that of stressed syllable /tai4/ is clearly prominent than those of unstressed syllables /bu4/, /yi4/, /da4/ and /li4/ respectively in this utterance.

Table 3: *Illustration of stress-induced pitch accent*

PW / Pitch level / Speaker	罗尔斯的	恢宏	巨著	正义论
F1	259.6	**276.4**	253.7	256.8
F2	188.7	**197.9**	173.5	194.4
M1	166.9	**177.9**	132.9	146.4
M2	117.5	**132.5**	96.3	110.9

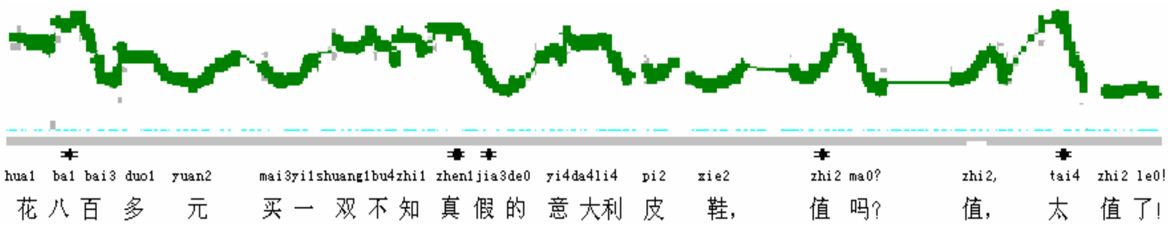

Fig.1: *Illustration of stress-induced pitch accent*

3.2 Physiological data and analyses

Due to the limitation of the corpus in original designing, in this part, the effect of boundary position was examined in prosodic phrase level, and the influence of stress contrast upon articulatory strengthening was tested in prosodic word level only.

Here, we just take the prosodic phrase of "说着说着"/shuo1zhe0 shuo1zhe0/ as an example, in which, both "说着" is a word contains one syllable with normal stress and followed by a weak-stress syllable, but the first "说着" occurs in the domain initial, while the second one in the domain final. Therefore, we can examine not only the effect of prosodic position-induced strengthening, but also the effect of stress-induced ones.

The preliminary EPG data measured from two speakers' (one male and one female) utterance were summarized in Table 4, and the strengthening effect was illustrated in Fig. 2. From these data, some interesting phenomena could be specified as follows.

3.2.1Articulatory strengthening at edges of prosodic domains

So far, from the limited measurements obtained here, an elongation of physiological duration has been discovered in the segments at edges of prosodic domain, it can be seen from the corresponding figures shown in Table 4, where both the data of male and female show that, (1) the physiological duration of consonant /sh/1 in first "说" is significantly longer than that of the consonants in all other positions within the phrase, it is a signal of longer closure interval and stronger energy holding during the articulation of this consonant. This fact gives an evidence for post-boundary strengthening of consonant articulation in phrase level. (2) The duration of /e/2 in second "着" is significantly longer than that of the vowels in all other positions within the phrase, it provides a strong evidence for pre-boundary strengthening of vowel articulation in phrase level. (3) The data referred to DRCS (duration ratio of closure vs. syllable or vowel vs. syllable) show an obviously longer closure in /sh/ 1 than that in all other positions' consonants, it reveals that post-boundary initial consonant does have more robust articulation; and at the same time, a definitely greater duration ratio is occupied by the /e/2 in second "着" within the phrase, this also confirms a pre-boundary articulatory strengthening. (4) According to the data of RCA (biggest ratio of linguo-palatal contact area), there seems no distinction on linguo-palatal contact area between /sh/ 1 and the consonants in other positions. On the other hand, however, the data does show a significant distinction between /e/1 and other vowels.

The phenomenon found in (4) above may be explained that (1)certain linguo-palatal contact is the necessity to form the articulation of a consonant, thus the contact must be kept relatively stable in contexts; (2)the amount of contact area is consonant-specified dependent on their different articulatory place. Consequently, the effect of prosodic positions on the consonant articulation can only manifested through longer contact duration, in stead of greater or less contact area. On the contrary, in the case of vowel, theoretically there should have no linguo-palatal contact during its articulation, though there do occur more or less contact here in real speech. Obviously, it is due to the contextual coarticulation. Therefore, the smaller RAC found for the /e/2 than other vowels can be explained that, the /e/2 has more typical or complete articulation and even overshot, while the others are more or less exposed by the coarticulation and obviously reduced or undershot.

In addition, the size of articulatory strengthening is also cumulative with the prosodic level. For instance, both /sh/1 in first "说" and /sh/2 in second "说" locates at the beginning of a word, but the strengthening of /sh/1 is significantly stronger than that of /sh/2. Definitely, it is because that /sh/1 is at the beginning of prosodic phrase, while /sh/2 at the beginning of prosodic word within the phrase. This situation is similar to that of acoustic manifestation described in 3.1, and the distance test result does show a strong correlation (R=0.98) between acoustic and physiological data.

From the picture shown in Fig. 2, the articulatory strengthening generally result in a greater phonetic distinction between the segments at boundary vicinity, that is, the feature contrast between consonant and vowel is more highlight. This fact means that hierarchical information of prosodic structure is exactly provided by both sides of the boundaries.

3.2.2 Articulatory strengthening at stressed locations

To examine the influence of stress status on articulatory strengthening, let's look at the data of the first /shuo1zhe0/ in Table 4 or Fig.2, a much stronger articulation of /shuo1/ can be discovered than /zhe0/ in terms of their physiological duration. Obviously, it is result in the stress contrast between the two syllables, and combined with the additional influence coming from the boundary position effect, here the stress-induced strengthening becomes more highlighted.

On the other hand, however, the situation in the second /shuo1zhe0/ seems more complex, i. e., the articulation of /sh/2 is stronger than that of /zh/2; while that of /uo/2 is weaker than that of /e/2 in terms of the DRCS data. It reflects that here the influence of stress contrast on the articulation is greatly restricted by pre-boundary effect.

In addition, /e/2 is absolutely longer than all other vowels in the phrase, though /e/2 is in a syllable with a weaker word stress. This is a signal that articulation of segments in prosodically strong positions are definitely strengthened even they are in phonologically weaker syllables.

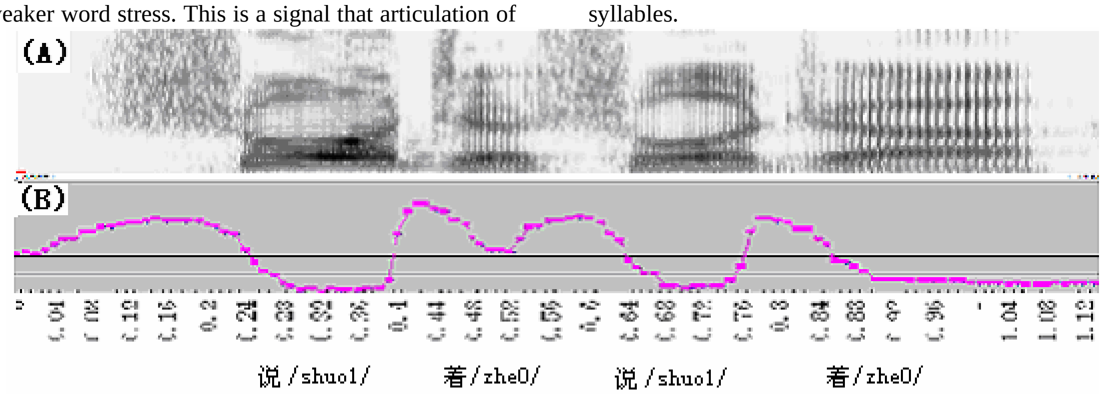

Fig.2: *(A) Spectrogram and (B)corresponding dynamic track of linguo-palatal contact during the articulation of the prosodic phrase of /shuo1zhe0 shuo1zhe0/ (“说着说着”)*

Table 5: *EPG measurements for the prosodic phrase of /shuo1zhe0 shuo1zhe0/ (“说着 说着”)*

Parameters \ speakers \ Segments		/sh/ 1	/uo/ 1	/sh/ 2	/uo/ 2	/zh/ 1	/e/ 1	/zh/ 2	/e/ 2
Duration (ms)	M	239.38	158.5	127.98	129.7	87.5	74.1	107.78	217.0
	F	740.02	187.13	180.79	150.65	128.45	87.2	144.31	244.22
RCA(%)	M	37	2	41	4	48	26	40	8
	F	50	7	47	6	46	37	46	3
DRCS(%)	M	43	57	24	76	41	59	15	85
	F	49	51	24	76	41	59	15	85

4. Conclusion

In summary, both acoustic and physiological data obtained from Mandarin Chinese reveal that phonetic manifestation of segments in connected speech is definitely moderated by their position or status in prosodic hierarchy, and it is evidenced by either segmental and suprasegmental parameters.

The patterns of strengthening are distinguished among other depending on their inducement: pre-boundary strengthening is dominant by syllable rhyme; post-boundary one is dominant by initial consonant; while stress / accent- induced one shows a relatively balanced distribution between initial consonant and final rhyme. Obviously, such distinction is well matched with their prosodic functions. Consequently, in certain degree, such patterns' different can serve as the markers of prosodic structure in spoken language processing.

In addition, according to the relationship discovered in 3.2.2, we would suggest that the influence from prosodic position is more powerful than that from stress contrast, and they seems to be integrated with a relation of algebraic sum.

All these effects observed here makes the picture more clear that articulators' movement during speaking is well-planned, such mechanism has been found in other languages, and now has confirmed here in Chinese.

References

[1] Meynadier, Y., 2004. Gradient linguopalatal variations due to a 4-level prosodic hierarchy in French. *9th Conference on Laboratory Phonology*, Urbana-Champain, IL, USA, June 2004.

[2]Cho, Taehong (2004) Prosodically-conditioned strengthening and vowel-to-vowel coarticulation in English. Journal of Phonetics 32 (2).

[3] Cao, J., 1999. Acoustic-phonetic characteristics on the rhythm of Standard Chinese, *Proc. of 4th National Conference on Modern Phonetics*, Beijing, August 25-27.

[4]Cao, J., 2004. Restudy of segmental lengthening in Mandarin Chinese. Proceedings of Speech Prosody'2004, Nara, Japan, 2004.

[5] Keating, P. A., 2004. Phonetic encoding of prosodic structure. *UCLA Working Papers I Phonetics, No. 103*.

[6] Cao, J.,2003. Rhythm of spoken Chinese: linguistic and paralinguistic evidence. *Proceedings of ICSLP'2000*, Beijing, Oct. 16-20, 2000.

[7]www.cass.net.cn/chinese/s18_yys/yuyin/product/product_6.htm

[8] Zheng, Y. & Zhu, S., 2001. The Dynamic EPG Corpus of Mandarin Speech and the Flat for Corresponding Research. *Acoustics and Electronic Engineering*, pp. 3-13.

[9] Campbell, N., 1993. Automatic detection of prosodic boundaries in speech, *Speech Communication* 13.