

Restudy of segmental lengthening in Mandarin Chinese

Cao Jianfen

Institute of Linguistics

Chinese Academy of Social Sciences

5 Jianguomennei Rd. 100732

Beijing, CHINA

Email: jianfencao@hotmail.com

ABSTRACT

This paper studies different types of segmental lengthening in Mandarin Chinese by clarifying temporal structure of various lengthened syllables. The discussion is mainly based on an investigation to a discourse speech corpus. The results show that there exist three types of syllable lengthening in spoken Chinese, and each type is matched with certain prosodic events and characterized by certain profile of temporal variation. Therefore, segmental lengthening may be a valuable source of information in understanding prosodic organization of speech. It should be of benefit to recognize the mechanism of speech production and perception, as well as improving in speech recognition, synthesis and natural language understanding.

1. Introduction

It is well-known that segmental lengthening can mark various prosodic events, such as phrase boundary and speech prominence. As one of the important cues of prosodic boundary, the domain-final syllable lengthening has been well studied. Besides, there are at least two more types of syllable lengthening in spoken language. The one is referred to speech prominence, and its lengthening profile is different from that of domain-final lengthening [1]; Another one is referred to the syllable of domain-initial position, but few attention was paid to this type of lengthening phenomenon.

Segmental lengthening is usually described in terms of syllable duration. In the fact, however, it is hard to distinguish different type of segmental lengthening if simply observe the global syllable duration. Our recent study has conducted a further analysis to lengthening phenomenon by clarifying the temporal variations within lengthened syllables in Mandarin Chinese. The present paper will discuss this aspect. The discussion is mainly based on an investigation to a discourse speech corpus, which is loudly read by multi-speakers. The results show that there exist different types of syllable lengthening which is characterized as prosodic position-dependent and stress status-dependent temporal variations, and this feature is highly identified across different speakers' speech.

2. Segmental lengthening at prosodic boundaries

As one of the functional markers in prosodic phrasing, syllable lengthening in domain-final position is most

prominent. It has been commonly agreed in phonetics area and applied in speech engineering. At the same time, our recent investigation has found that the syllable in domain-initial position exhibit obvious lengthening, too, and the pattern of domain-initial (hereafter domain-start) lengthening is obviously different from that of domain-final (hereafter domain-end) ones.

2.1 Segmental lengthening in prosodic domain-end position

Syllable lengthening in prosodic domain-end position usually to be called pre-boundary lengthening or unit-final lengthening.

It has been commonly agreed that there does exist pre-boundary lengthening in spoken Chinese; however, some disagreements were raised from different studies. For example,

according to an investigation to the news speech and prose declaim, Cao[2] suggested that pre-boundary lengthening mainly exists in prosodic phrase level, but generally not in sentence and paragraph level. While Feng's approach [3] argued that it also occurs in sentence level, and the difference of lengthened degree between phrase and sentence level is not significant though there is some difference there.

Table 1 *Average duration (ms) of syllables in different prosodic positions*

Level Speaker		Sentence	Prosodic phrase	Prosodic word	Utter. mean	
F1	start	209.2	171.3	178.7	<i>Avg.</i>	178.7
	end	245.2	251.4	139.9	<i>Sd.</i>	139.9
F2	start	221.4	213.7	173.3	<i>Avg.</i>	173.3
	end	208.0	257.5	161.0	<i>Sd.</i>	161.0
M1	start	202.4	201.3	181.7	<i>Avg.</i>	181.7
	end	195.2	247.2	203.5	<i>Sd.</i>	203.5
M2	start	232.6	174.4	182.4	<i>Avg.</i>	182.4
	end	174.0	196.4	168.0	<i>Sd.</i>	168.0
Average	start	216.3	190.2	179.0	<i>Avg.</i>	179.0
	end	196.6	235.5	168.1	<i>Sd.</i>	168.1

To clarify this issue, an investigation to a speech corpus of read discourse was conducted. From the results summarized in Table 1, we can observed that a domain-end lengthening occurs significantly ($p=0.006$) only in prosodic phrase level, but slightly and occasionally in sentence level and it is not significant ($p=0.668$) if compared with general mean in the utterance. In addition, from the bold figures

shown in Table 2 and the relevant figures in Table 3, which were obtained from a further study referred to temporal distribution within the syllables, we can see that the domain-end lengthening occurred in prosodic phrase level is contributed only by syllable-final (rhyme), and the syllable-initial (consonant) is even shortened in that case. While in sentence level, the lengthening, if there is any, is mainly manifested by syllable-initial but occasionally by syllable-final, and it is speaker-dependent, in some case, the syllable-final is even shortened. Moreover, such distinction existed between the two levels is systematical and identified across different speakers.

Consequently, if taking syllable-rhyme duration as the judgement criterion, then the domain-end lengthening is limited only in prosodic phrase level in spoken Chinese.

Table 2 *Average duration (ms) of syllable-final in different prosodic positions*

Level Speaker		Sentence	Prosodic phrase	Prosodic word	Utter. mean	
F1	start	144.4	126.5	122.5	Av.	143.8
	end	185.2	202.2	126.2	Sd	48.3
F2	start	162.4	138.7	129.4	Av.	149.4
	end	153.0	210.7	140.4	Sd	49.5
M1	start	127.4	145.5	121.2	Av.	143.0
	end	69.0	170.1	125.4	Sd	51.0
M2	start	127.8	116.5	93.9	Av.	126.7
	end	113.6	148.9	119.0	Sd	48.7
Average	start	140.5	131.8	119.1	Av.	141.0
	end	130.5	183.0	127.8	Sd	49.9

Table 3 *Average duration (ms) of syllable-initial in different prosodic positions*

Level Speaker		Sent.	PP	PW	Utter. mean	
F1	start	80.5	70.3	58.2	Av.	55.3
	end	60.0	52.7	39.6	Sd	27.1
F2	start	97.0	98.7	85.1	Av.	69.9
	end	76.0	62.5	39.2	Sd	36.8
M1	start	98.0	76.7	67.9	Av.	64.0
	end	70.0	59.1	41.1	Sd	33.0
M2	start	108.8	78.5	66.9	Av.	56.0
	end	61.0	50.3	34.3	Sd	31.1
Average	start	96.1	81.0	69.1	Av.	61.3
	end	66.8	56.1	38.6	Sd	32.6

2.2 Segmental lengthening in prosodic domain-start position

Similar to the case in domain-end position, syllables in domain-start position show a lengthening, too.

However, the lengthening pattern in domain-start position is quite different from that of domain-end ones. As can be seen from Table 3 and 2, the lengthening in prosodic domain-end position is manifested by syllable-final, while that in domain-start is manifested by syllable-initial consonant. If taking the duration ratio of syllable-initial to syllable-final as the scale and comparing with that of general mean (0.48) in utterance, then the ratios in the start position

of sentence, prosodic phrase (PP) and prosodic word (PW) are 0.68, 0.61 and 0.58 respectively. Obviously, syllable-initial lengthening does commonly exist in prosodic domain-start position no matter in which speech level. In contrast to the pre-boundary lengthening, domain-start lengthening may be called post-boundary lengthening.

In the fact, domain-start lengthening has its articulatory bases. Early around 1990's, several studies had shown that articulations of tongue, lips and velum were different in magnitude in domain-initial (start) position vs. non-initial position. For example, English alveolar and velar stops have more contact between the tongue and the palate when they are word-initial than word-final [4]; Fougeron, C. & P. Keating's work [5] also show that English consonants are articulatorily strengthened when they are initial in a prosodic domain. All these articulatory features mentioned above have been referred to so-called prosodic strengthening, thus, I would suggest that domain-start segmental lengthening observed here should be a corresponding result of articulatory strengthening.

3. Segmental lengthening in accented syllables

As one of the acoustic correlates, longer duration is usually observed from the accented syllable in speech, however, the specific situation seems to be language- and category-dependent.

Generally, there is two main categories of sentence stress in Chinese: grammatical (i.e., default) stress and logical (i.e., marked) stress. Our approach examined syllable duration in these two categories respectively.

3.1 Segmental lengthening in the syllables with default (grammatical) stress

At first, we examined the situation in the default case. The measured data is listed briefly in Table 4, where the data of stressed syllables were measured from all of non-terminal prosodic words in the utterance. Obviously, in Chinese, according to the data shown in Table 4, there seems no lengthening phenomenon referred to default stress, no matter it is compared with the mean value in non-terminal prosodic words or in the whole utterance.

Table 4 *Average duration (ms) of syllable and duration distribution within the syllable in default stress*

Condition Distribution	in default stressed words	in non-terminal prosodic words	in whole utterance
Syllable	192.8	195.5	192.1
Syll-initial (consonant)	55.7	55.6	61.3
Syll-final (rhyme)	135.7	142.6	140.8
Ratio of initial vs. syllable	0.289	0.284	0.284
Ratio of final vs. syllable	0.704	0.704	0.733

3.2 Segmental lengthening in the syllables with marked stress

Table 5 *Average duration of syllables in the sentences with different logical stress*

Context syllables		In sent. 1	In sent. 2	In sent. 3	In sent. 4
Zhe	ms	190	200	280	230
	ratio	14.5	13.8	21.4	15.3
Shuang	ms	220	240	240	260
	ratio	16.8	16.6	18.3	17.3
Xie	ms	230	250	210	380
	ratio	17.6	17.2	16.0	25.3
Bu	ms	110	270	100	80
	ratio	8.4	18.6	7.6	5.3
Jie	ms	330	310	280	320
	ratio	25.2	21.4	21.4	21.3
Shi	ms	230	180	200	230
	ratio	17.6	12.4	15.3	15.3

Comparing with default stressed syllable, the situation for the case of marked stressed ones, i.e., the logical (including contrastive, emphasized) stressed syllable is quite different. We examined the temporal distribution to a set of sentences that consists of the same syllable string but with different logical stress assignment, these materials were read by 9 speakers. The general results were summarized in Table 5, where the bold figures show the duration value of logically stressed syllable in each sentence. If comparing the bold figures with their partners in different sentences, an obvious syllable lengthening can be observed definitely. In addition, this phenomenon is existed commonly across all the 9 speaker.

4. Comparison among different segmental lengthening

Table 6 summarizes the durational distribution in different category of lengthened syllables. From the bold figures shown in the table, the distinction of lengthening profile between pre- and post-boundary cases is confirmed again, and that of the syllable with marked stress seems to be more related to its initial consonant, it is similar to but slightly weaker than that of post-boundary lengthening. Therefore, if compare with that of boundary motivated lengthening, the lengthening profile of marked stress is

relatively balanced within a syllable.

Table 6 *Comparison between the lengthening of pre- and post-boundary as well as marked stress case in terms of duration distribution within a syllable*

Conditions	Duration ratio of initial to syllable.	Duration ratio of final to syllable
In sent. start	0.444	0.649
In PPstart	0.425	0.693
In PWstart	0.386	0.666
In sent. end	0.309	0.603
In PPend	0.294	0.962
In PWend	0.216	0.719
In marked stress case	0.369	0.631
In marked stress case	0.319	0.733

5. Conclusion

According to the investigation described above, segmental lengthening may be classified into three categories, namely, pre-boundary, post-boundary and prominence related ones. Each of them is characterized by syllable-final dominant, syllable-initial consonant dominant and syllable-initial and -final balanced respectively. Obviously, the distinction between them is well matched with their prosodic functions. Consequently, it can serve as the markers of prosodic segmentation and prominence allocation in spoken language processing.

Reference

- [1]Campbell, N., 1993, Automatic detection of prosodic boundaries in speech, *Speech Communication* 13.
- [2]Cao, Jianfen, 1998, A preliminary study of rhythm in Mandarin Chinese. *RPR-IL(CASS)/1998*.
- [3]Feng, Yongqiang et al. , 2001, A statistical analysis to syllable duration of Chinese. In *Proc. of 5th National Conference on Modern Phonetics*, Beijing.
- [4]Byrd, D. M., 1994, Articulatory timing in English consonant sequences. PhD. Diss. *UCLA Working Paper in Phonetics*, 86.
- [5]Fougeron, C. & P. Keating, 1996, Articulatory strengthening in prosodic domain-initial position. *UCLA Working Paper in Phonetics*, 92.