Restudy of segmental lengthening in Mandarin Chinese

CAO Jianfen

jianfencao@hotmail.com

Institute of Linguistics, Chinese Academy of Social Sciences

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 (Campbell, 1993); 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 (1998) suggested that pre-boundary lengthening mainly exists in prosodic phrase level, but generally not in sentence and paragraph level. While Feng's approach (2001) 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	Av.	199.3
	end	245.2	251.4	139.9	Sd	35.9
F2	start	221.4	213.7	173.3	Av.	205.8
	end	208.0	257.5	161.0	Sd	25.7
M1	start	202.4	201.3	181.7	Av.	205.2
	end	195.2	247.2	203.5	Sd	13.9
M2	start	232.6	174.4	182.4	Av.	187.9
	end	174.0	196.4	168.0	Sd	17.6
Average	start	216.3	190.2	179.0	Av.	197.6
	end	196.6	235.5	168.1	Sd	18.8

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

---

[*] This paper has been published in *Proceedings of Speech Prosody'2004*, Nara, Japan, 2004.

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 judgment criterion, then the domain-end lengthening is limited only in prosodic phrase level in spoken Chinese.

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 mater 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 (Byrd, 1994); Fougeron, C. & P. Keating's work (Fougeron, 1996) also shows 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 are 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 speakers.

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 PP start	**0.425**	0.693
In PW start	**0.386**	0.666
In sent. end	0.309	0.603
In PP end	0.294	**0.962**
In PW end	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.

References

Byrd, D. M., 1994. Articulatory timing in English consonant sequences. PhD. Diss. UCLA Working Paper in Phonetics, 86.

Campbell, N., 1993. Automatic detection of prosodic boundaries in speech, Speech Communication 13.

Cao, Jianfen, 1998. A preliminary study of rhythm in Mandarin Chinese. RPR-IL(CASS)/1998.

Feng, Yongqiang et al. , 2001. A statistical analysis to syllable duration of Chinese. In Proc. of 5th National Conference on Modern Phonetics, Beijing.

Fougeron, C. & P. Keating, 1996. Articulatory strengthening in prosodic domain-initial position. UCLA Working Paper in Phonetics, 92.

汉语普通话音段的韵律延长效应再探

曹剑芬

中国社会科学院语言研究所

摘要[*]

本文通过分析音节的时长结构，研究各种韵律延长的类型及其不同特点。研究和讨论的基础主要是对连续话语语料库语音的测量和分析。结果发现，在汉语口语里，存在着三种不同类型的音段延长，每一种类型都与一定的韵律事件相对应，都有各自独特的时域变化方式。具体说来，边界前的音段延长以音节韵母的延长为主，边界后音段延长以声母辅音为主，而与重音凸显相关的音段延长则是涉及整个音节的、相对平衡的延长。因此，音段的不同延长效应，也许可以作为了解言语韵律结构的有价值的信息。这对于加深对言语产生及感知机制的认识、提高语音合成和语音识别以及自然语言理解的效率无疑是非常有益的。

Tonal Aspects in Spoken Chinese: Global and Local Perspectives

CAO Jianfen
jianfencao@hotmail.com

Institute of Linguistics, Chinese Academy of Social Sciences

Abstract[*]

The present study try to clarify the deep essence and structure of Chinese intonation by objectively approach to the production process of tone and intonation and their interaction based on the acoustic-phonetic observation. Our main attention is paid to examine (1) what is the basic intonation pattern for the whole course of utterance and how it is formed? (2) How the lexical tone patterns integrate with the global intonation pattern? The preliminary results indicate that intonation pattern is mainly related to pitch register movement of global utterance; it is undulated depending on physiological mechanism and the needs of speech mood and semantic expression. While tone pattern is mainly referred both to the contour and the register of pitch change of local syllable or word, and it is lexically given. In real speech, each tone must be modified by global intonation through adjusting its relative register on one hand, and keeps its basic tone shape on the other hand. At the same time, the global intonation must be manifested through the pitch movement of each local tone. The key point is that the relationship between tone and intonation is an "algebraic sum" of pitch register, in stead of that of pitch contour.

1. Introduction

Intonation is the melodic pattern of a language, objectively referred to the pitch movement of an utterance, while its manifestation is language specific.

In respect to spoken Chinese, tonal aspects are rather significant in its function and particularly complex in its variation. Functionally, tonal variation can signal lexical identity, speech prominence, rhythmic organization, as well as intonation structure. These aspects are synchronically carried through the same entity of fundamental frequency, but behaved in different way and realized as different prosodic component in natural speech, thus, result in a particularly complex F0 contour in surface. Consequently, how to expose the relationship among these components becomes an urgent issue in phonetics theory, and how to decompose them from the surface F0 contour is one of the most difficult tasks in spoken Chinese processing.

To explore the work mechanism for each aspect of prosody in Chinese, many contributes have been made, and various controversies were raised. Among these challenges, the most prominent issue was concentrated on the relationship between lexical tone and intonation. The key point is that how the local tones integrate with global intonation? In the other words, how the individual syllable or word to be said in any intonation without losing its lexical identity? In respect to this issue, various theories have raised since early in last century. The most famous one is the so-called "small ripples riding on large waves" theory as suggested by Chao (1958, 1980). Unfortunately, however, this theory is still remaining as controversial (Shen, 1985, 1992), the focus of arguments is that how to view the "algebraic sum" in that theory. Shen (1992) argued that the relationship between tone and intonation seems difficult to be counted as an "algebraic sum". While Wu (1996) interpreted that

Chao's "small ripples" to be elevated by "large waves" is result in the increasing of tonal scale, but not changing of its contour shape, so the relation of "algebraic sum" is not difficult to be counted.

Another related issue is that whether there is an explicit global form for the intonation of an utterance? It was raised recently by Xu (1999, 2001). After a series of profound approaches, he suggested that there is no explicit global form for the intonation of an utterance. The surface F0 declination is determined by the role of multiple sources, such as the downstep caused by L tone, the new topic or focus in the utterance, and these effects are parallel.

However, according to Shih's (2001) experiment, a clear declination effect was observed. She investigated to a set of sentences, which were designed consist of syllables all with high level tones, so it has avoided downstep effect from low tone features.

Recently, Wang (2003) investigated to 600 sentences which was designed in consists of four lexical tone sequences respectively. Her findings confirmed that there does exist an underlying F0 declination besides the downstep effects come from the low tone feature, the new topic and focus.

However, both of Shih and Wang's research were based on designed read sentences, and these sentences are all consists of syllables with single category of tones. Whether their findings are also true in natural speech or not, some further study is needed.

The present paper will discuss these aspects based on our recent investigation. The main attention is paid to examine (1) what is the basic intonation skeleton for the whole course of utterance and how it is formed? (2) How the lexical tones integrate with the global intonation on one side, and keep their own patterns respectively as well? The preliminary results obtained from this investigation may lead to a practical model of intonation for text-to-speech synthesis in Mandarin Chinese.

2. Material and method

The material used in this study is a set of discourse read by 4 different speakers, through which, we try to discover if there is any global frame of intonation in real Chinese speech.

Acoustically, both tone and intonation is time-varying pattern of pitch change, and both of them are referred to the movement of relative pitch register as well. Hence the pitch value for each syllable was measured, and corresponding register was calculated.

3. Results and Discussion

3.1 Results

Preliminary results are summarized in Table 1, where the distribution of pitch register of syllables in terms of average pitch value in the utterance is listed. These data were calculated according to the upper and lower pitch value of each syllable in certain position of 4 speaker's speech.

---

[*] This paper has been published in *Proceedings of International Symposium on Tonal Aspects of Languages: Emphasis on Tone Languages*, Beijing, March 28-30, 2004-6-28.

Table 1: Average pitch value of syllables with different tone categories at different positions in the utterance

Condition / Position	Mean value	1st tone	2nd tone	3rd tone	4th tone	Neutral tone
Sent. start	196.2	224.4	182.3	154.8	223.1	--
Sent. mid	164.7	182.8	158.1	149.1	171.2	162.3
Sent. end	113.6	107.0	115.7	--	118.1	--
Phrase start	184.4	190.2	182.6	159.1	205.8	--
Phrase mid	164.7	182.8	158.1	149.1	171.2	162.3
Phrase end	143.7	143.4	152.0	--	154.9	124.6
Gen. mean	166.7	180.1	161.7	152.5	169.2	154.0

For the convenience of description, a diagram is given in Fig.1.

shi4jian1de0   yuan2lai2 ji2   ke3yi3yong4 ji2   jia1yi3
wen4ti2   fu4za2de0   jian3dan1de0 shi4li4   shuo1ming2
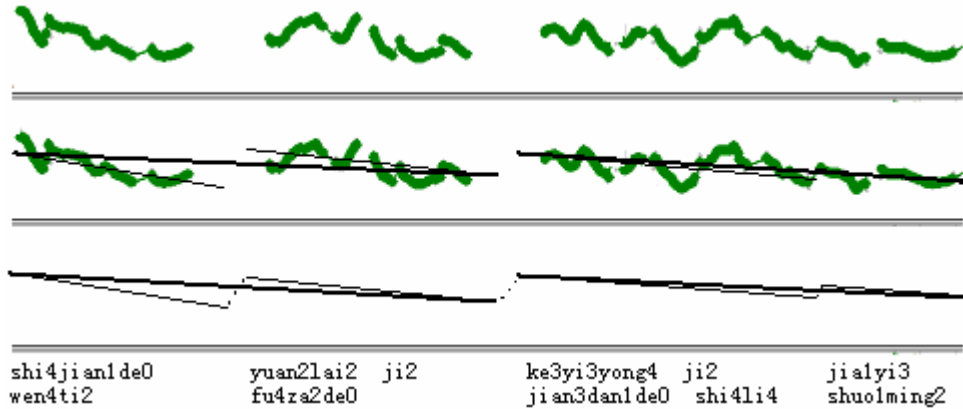
Fig. 1: Example of pitch movement of an utterance

3.2 Discussion

3.2.1 Global pattern (basic skeleton) of pitch movement in phrase and sentence

First, from the data listed in Table1, a declined trend of pitch register can be seen clearly both in phrase and sentence layer, and it is systematically identified no matter in general situation or in the case of different tone categories. That is, in all the cases, the pitch register at sentence start is higher successively than that in sentence mid and end, and so is in phrase layer. Actually, due to the physiological constraint during speech production, a gradually declination trend of F0 movement must be taken place. Because during speaking, the effect of dropping in speaker sub glottal pressure upon the F0 declination is a kind of articulatory constraint naturally taken place in speech production, and it needs not to be pre-planned. The declination phenomenon observed here indicates that how robust the declination mechanism is in real speech, and it is right the base of intonation skeleton.

Second, such declination is hierarchically organized as shown in Fig.1. Here the picture in top line of the Fig is the original pitch curve; the ones in mid is a diagram of declination trend, where several declined lines added on the original pitch curve, they represent the moving direction of pitch register, they were drawn according to average pitch value of first syllable and last syllable respectively for each prosodic unit. For the convenience of observation, they were separated out as the broken lines at the Fig's bottom. From this Fig., the hierarchical structure can be more directly observed. On one hand, the declination trend is formed in sentence and phrase level respectively, and it is relatively independent in each level, thus form as different layers. On the other hand, the declination profile of sentence pregnancies with those of phrases, thus form into an organic unity. In the fact, the size of phrase is usually different and the smaller ones are embedded in the larger ones. Consequently, the declination profile of an utterance has more complex hierarchy. In addition, as can be seen from the bottom line in Fig. 1, these declined trend in different layers are reset constantly in

corresponding prosodic boundaries, thereby leads to the undulation of intonation "large wave" of the utterance. Comparing to the pitch movement within a syllable, the pitch undulation referred to intonation is relatively slow and simple. Moreover, the specific declined slope in speech will be further modified by speech mood, focus distribution and rhythmic structure, these modifications are also achieved through pitch register, and the details will be specified in another paper (Cao, 2004).

3.2.2 Local pattern of pitch movement in syllable or word

The pitch movement of a syllable or word in running speech usually contain two aspects: the one is the time-varying pitch contour (shape) of the syllable, it is determined by its tonal pattern including lexical tone sandhi rule (1990); the other is the relative pitch register of the syllable or word, it is phonologically constrained and must be modified by rhythm, stress and intonation of the utterance.

In the most cases, such lexically given tone patterns are 1 to 3 syllabic groups. From Fig. 1, such tone patterns can be seen everywhere over the utterance, thus forms the so-called "small ripples". Of course, however, the tone patterns in the utterance are somehow deviated from their original form in isolation, which will be specified in section 3.2.3.

3.2.3 Integration between local tones and global intonation

(1) Exist simultaneously

According to Fig.1, we can see that in continuous speech, the local tones keep their basic F0 pattern, but aligned follow the declined trend of global intonation by elevating or lowering their register. For example, in the first sentence, syllables "shi4" and "wen4" belong to the same lexical tone category, they have almost the same contour pattern, but different in pitch register: the former is obviously higher than the later. Similarly, syllables "yuan2" and "za2" in second phrase, and the two "yi3" in second sentence, etc., each pair of them have similar pitch contour, but

different in their register, the former ones' are all higher than that of the later ones. All these phenomena indicate that the information of tone and intonation are carried synchronically by the same entity.

(2) Superposed each other as algebraic sum

In the case that, if a lexical unit, which is characterized by a relatively higher pitch register, for example, that with a 1st tone or 4th tone, like the "shi4", occurs nearby the peak of "large waves", its pitch register is further elevated. If a syllable characterized by high register, like the "wen4" occurs nearby the valley of "large waves", its register is relatively lowered. On the contrary, if a lexical unit is characterized by low register feature, for example, the "yi3" with a 3rd tone which is characterized by low register, when it occurs nearby the peak of "large waves", its pitch register is relatively elevated; while when it occurs nearby the valley of "large waves", its register is further lowered. That is right agreed with Chao's(1980) opinion of "positive plus positive is more positive" and "positive plus minus will be dependent on which one's absolute value is larger". Accordingly, the key point of so-called "algebraic sum" between local tone and global intonation is mainly related to the movement of pitch register, in stead of the variation of pitch contour.

4. Conclusion

According to the preliminary results obtained from this investigation, we would suggest that (1) there does contain an explicit form of entire intonation component, it presents as the declined trend of pitch register. This trend is slowly downward and upward alternately due to the resetting at different prosodic boundary. We suppose that such underlying declination is mainly caused by physiological constraint, and strengthened by the downstep effect of L tone, focus and new topic. The specific declined slope in real speech will depend on speech mood, focus distribution and rhythmic structure. (2) Our data reveal that Chao's "algebraic sum" theory is basically reliable, but his attention was mainly concentrated on the last syllable of a phrase or sentence, while ignore other parts of the utterance. Our data support this theory and further clarify that the relationship of "algebraic sum" between tone and intonation does work over the whole course of the utterance. (3)Local tones and global intonation in Chinese is both relatively independent and integrate with each other. On the one hand, the global form of intonation is manifested through local configurations of tones; on the other hand, each local configuration of syllables must be modified by global configuration, thus synchronically carries the information of lexical tones and intonation. Their integration is achieved through the variation of pitch register, instead of pitch contour.

References

Cao, Jianfen, 2004, Intonation structure of spoken Chinese: its universality and specificity. *From Traditional Phonology to Modern Speech Processing* (*Festschrifts for Celebration to Professor Wu Zongji's 95th Birthday)*, Foreign Language Teaching and Research Press, Beijing.

Chao, Yuan-ren, 1958-1980, *The problem of language.* Commercial Press, Beijing.

Shen, Jiong, 1985, Pitch range of tone and intonation in Beijing dialect. In Lin Tao et at. Edited *Experimental Phonetic Study of Beijing Dialect*, Beijing University Press.

Shen, Jiong, 1992, An intonation model of Chinese. *Yunwen Yanjiu,* No.4.

Shih, Chilin, 2001, Generation and normalization of tonal variations. *Journal of Chinese Linguistics*, monograph series, no. 17, University of California, Berkeley, USA.

Wang, Anhong, 2003, *Research on the pitch downtrend of intonation in Putonghua*, The PhD dissertation of Beijing University.

Wu, Zong-ji, 1990, Can polysyllabic tone-sandhi patterns be the invariant units of intonation in spoken Standard Chinese? *Proc. of ICSLP'1990*, Kobe, Japan.

Wu, Zong-ji, 1996, Yuan-ren Chao's contribution on the tonal study of Chinese. *Journal of Tsinghua University*, Vol. 11, No. 3.

Xu, Yi, 1999, Effects of tone and focus on the formation and alignment of F0 contours. *J. Phonetics* 27, 55-105.

Xu, Yi, 2001. Sources of tonal variations in connected speech, Journal of Chinese Linguistics, monograph series #17:1-31.

汉语口语的声调和语调—总体特性和局部特性

曹剑芬

中国社会科学院语言研究所

摘要[*]

本文以声学-语音学的实验分析为基础，力图探索汉语声调和语调生成的过程，以便客观地揭示汉语语调的深层本质及内部结构。我们的注意力主要集中在以下两个问题上：（1）什么是话语语调的基本模式？它是怎样形成的？（2）具有区别词义功能的声调是怎样跟话语的总体语调有机结合的？初步的考察结果表明，话语语调的基本模式主要跟话语的总体音域、即调阶运动有关，它的起伏波动取决于言语产生的生理制约以及语气、情感及语义表达的需要。而声调模式既跟各个局部音节或词的音高变化的曲拱调形相关，又跟它们音高变化的音阶特征有关。在实际话语中，各个局部的音节或词，一方面基本保持它们的音高曲拱模式，以满足词义区别的需要，另一方面通过它们的音阶的相对起落变化，满足话语总体语调调阶起伏的需要。也就是说，在实际话语里，局部的声调跟总体的语调是通过它们的音阶的相互叠加而有机结合起来的。