

# Tonal Aspects in Spoken Chinese: Global and Local Perspectives

CAO Jianfen

[jianfencao@hotmail.com](mailto:jianfencao@hotmail.com)

Institute of Linguistics, Chinese Academy of Social Sciences

## Abstract<sup>\*</sup>

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.

<sup>\*</sup> 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	1 <sup>st</sup> tone	2 <sup>nd</sup> tone	3 <sup>rd</sup> tone	4 <sup>th</sup> 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.

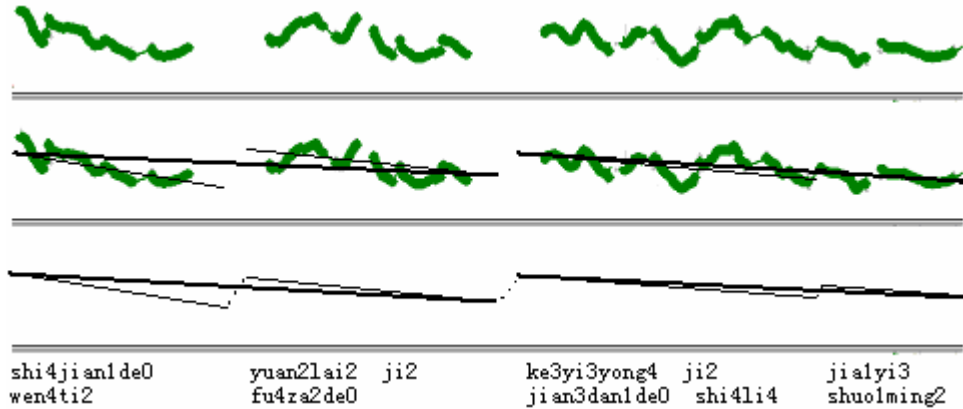


Fig. 1: Example of pitch movement of an utterance

## 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 1<sup>st</sup> tone or 4<sup>th</sup> 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 3<sup>rd</sup> 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 95<sup>th</sup> 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 al. 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阶、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。

\* 原载《国际声调和语调研讨会论文集》,北京,2004年3月28-30日。