

Tone Pattern and Word Stress in Mandarin

Wang Yunjia

Department of Chinese Language and Literature

Peking University, Beijing

wyj@chinese.pku.edu.cn

Abstract

The relationship between tone pattern and word stress is investigated on different aspects. The distribution of the original tones of the unstressed syllables in disyllabic words, the perception of word stress in utterances and the allocation of sentence stresses within words are analyzed. The results of the analyses coincide quite well, indicating that tone has something to do with the engendering of the unstressed syllables at lexical level and with word stress pattern in continuous speech. Specifically, when appearing in the final position in a disyllabic word, the high level tone, i.e. T1, makes a syllable have the smallest possibility to be unstressed at lexical level, and also makes it most prominent in perception, and the low tone, i.e. T3, does the reverse. It is argued from the diachronic viewpoint that to some degree the prominence of a tone in continuous speech makes it uneasy to become neutralized at lexical level.

1. Introduction

The problem that if Mandarin, which is a tonal language, has word stress has been discussed for a long time, and up to now researchers do not reached any agreement. Nevertheless, people agree that there are two stress categories at least at lexical level, i.e. normal stress that is loaded by a syllable with a normal tone and weak stress that is loaded by a syllable with the neutral tone. Chao argued that in a polysyllabic phrase or word that does not contain any neutralized syllable, the stress degrees of the syllables are not equal [1], and this claim is also widely accepted by phoneticians. Since disyllabic words are in the majority of the words in modern Mandarin, they are usually the focus in the studies on word stress pattern. Providing that word stress in Mandarin is dualistic contrastive, disyllabic words can be divided into two groups, i.e. the words consisting of two normally stressed syllables (SS words) and the ones consisting of one normally stressed syllable and one unstressed syllable (SU words).

In regard to word stress in Mandarin, people pay more attention to the unstressed syllables than they do to the stressed ones. Under most circumstances, the distribution of SU words can be predicted from lexical construction. For instance, most of the suffixes are unstressed. However, there are some SU words within which tone neutralization of syllables has nothing to do with lexical construction [2], e.g., “聪明”, “闺女”, and “口袋”. It is widely said that these words are frequently used and have a long history in Mandarin. Nevertheless, some of the SU words are quite new, e.g. “干部”, “意见”, and only some of the frequently used and old words contain unstressed syllables. Therefore, it seems that there exists a factor beyond lexics that makes the tone of the final syllable neutralized in some disyllabic words.

It is well known that pitch is the closest acoustic correlate to word stress in English as well as some other stress

languages [3][4][5]. It was also proved by some experimental studies that pitch plays an important role in the perception of the unstressed syllables in Mandarin [6][7]. Since tone is the pitch pattern of a syllable, it is reasonable to ask from the diachronic viewpoint if the original tone pattern of a syllable was one of the factors that made the syllable unstressed in a SU word. Li argued that Tone 4 in Mandarin seems easier to be neutralized than the other tones do, since he found that, in *A Modern Chinese Dictionary* (version 1992) [8], the SU words within which the second syllables originally carried Tone 4 account for 36.64% of the total SU words. However, he pointed out that since the amount of the SS words ending by Tone 4 is larger than those of the words ending by the other tones, the proportion 36.64% couldn't prove his hypothesis very strongly [9]. Li's analysis of his data is rational; still, to conclude that his hypothesis is null and that tone has nothing to do with the engendering of unstressed word is unreasonable. What one should do further on this problem is to find a better statistical method to analyze the data.

Since a neutral tone always follows a normal tone in a disyllabic word, and the pitch of the neutral tone depends on the tone it follows [10], it is also reasonable to ask if the initial tone co-worked with the original tone of the unstressed syllable in the engendering of an SU word. Therefore, this paper tries to investigate whether tone neutralization of the final syllables in SU words has something to do with the original tonal combination of the words, and if so, in which kinds of tone combinations the final syllables have relatively larger possibility to be neutralized. The factors at the levels beyond phonetics that cause the engendering of the SU words will be excluded in the consideration of this study. To make the narration concise, the four normal tones in Mandarin will be abbreviated as T1, T2, T3, and T4, respectively.

2. The distribution of SU words in 16 tonal combinations

In this section we use the method similar to what Li used [9], namely, to infer the chronic variation from the synchronic material, but our material is different from Li's.

There are 16 kinds of tonal combinations of disyllabic words in Mandarin. We will analyze the distribution of SU words among 16 tonal combinations by observing the disyllabic words in *A Dictionary of the Frequencies of the General Words in Modern Chinese* [11]. 30078 SS words and 932 SU words are used in this study, the SU words in which the unstressed syllables are syntactical words (e.g. “的”, “了”, “着”) and the words in which the original tones of the unstressed syllables are unclear (e.g. “马虎”, “咕噜”, “什么”) being excluded.

The original tones of the unstressed syllables of the SU words are labeled, thus, the original tonal combination of each of the studied words falls into one of the 16 tonal

combinations, no matter whether a word is a SS one or a SU one. From the viewpoint of lexical construction, there are at least three groups of unstressed syllables here, the first group is suffix, e.g. “子(桌子)”, “当(顺当)”, “气(秀气)”, the second one is the reduplication of the initial syllable, e.g. “叔(叔叔)”, “爷(爷爷)”, “套(套套)”, and the third one is root, e.g. “思(意思)”, “白(明白)”, “指(戒指)”. There are 516 SU words in which the unstressed syllables are suffixes or the reduplications of the initial syllables and 416 US words in which the unstressed words are roots.

As we mentioned in the introduction, the ratio of the SU words to the total of the SU words is almost meaningless, so we have to use other parameters of the SU words. We first calculate the ratio of SU words to all the words for each of the 16 original tonal combinations. Considering that a few suffixes (e.g. “子”, “头”) appear in a larger proportion of the SU words, two methods are used in the calculation of the ratio, being expressed by equation (1) and (2).

$$R1 = \text{number of the SU words} / \text{sum of the SS words and the SU words} \quad (1)$$

$$R2 = \text{number of the SU words in which the unstressed syllables are root} / \text{sum of the SS words and the SU words} \quad (2)$$

For instance, for the original tonal combination of “T1+T1”, there are 1974 SS words, 44 SU words, and 24 SU words in which the unstressed syllables are roots, then $R1=44/(1974+44)=0.022$, and $R2=24/(1974+24)=0.012$.

Obviously, the larger a ratio for a tonal combination, the more easily an original SS word of this combination turns to a SU word, so what we should do secondly is to observe the relationship among the 16 ratios. Fig.1(a) and (b) show $R1$ and $R2$ for 16 tonal combinations, respectively, TIS in the figure being the abbreviation of the tone of the initial syllable and remaining so in the following part of this paper.

We can view the statistics in (a) and (b) at two points, i.e. the initial tone and the final tone. Looking into Fig.1(a) first, we cannot find out any regular relationship among the ratios under different initial-tone conditions. The relationship among the ratios under the four final-tone conditions, however, is quite clear, i.e. $R1/T3 > R1/T2 > R1/T4 > R1/T1$, no matter which tone the initial syllable carries. It is remarkable that the ratios of the SU words under the final-tone condition of T3 are much larger than those under the other final-tone conditions. In Fig.1(b), the relationship among the ratios under different initial-tone conditions is still irregular. The relationship among the ratios under different final-tone conditions has both similarity and dissimilarity to that in Fig.1(a). The similarity is that the ratios under the final-tone condition of T4 are larger than the ones under the final-tone condition of T1 in both (a) and (b). The dissimilarity is that unlike what it does in (a), $R2$ under the final-tone condition of T2 or T3 in (b) co-varies with the initial tone. When the final syllable bears T2, the ratio for the combination of “T3+T2” is quite small, while the other ratios are quite large. When the final syllable bears T3, the ratio for the combination of “T1+T3” is the largest one among all the ratios, and the other ratios are the smallest ones.

The discord between $R1$ and $R2$ is caused by the unbalanced distribution of the original tones of the neutralized suffixes. Among the studied words, the amounts of the suffixes originally bearing T1, T2, T3, and T4 are 9, 69, 389, and 17, respectively. Supposing that in suffix engendering,

meaning losing of a morpheme synchronized with or preceded tone neutralization, to some degree the numbers indicate that a suffix originally bearing T3 has the largest possibility to be neutralized, and that a suffix originally bearing T1 has the smallest, though the large amount of the suffixes originally bearing T3 is mainly owed to the fact that the suffix “子” is extraordinarily combinable in word construction.

In summary, the tone of the initial syllable does not show a significant effect on the engendering of SU words. Regardless of word construction, T3 makes a syllable to have the largest possibility to be unstressed. T4 makes a syllable to be unstressed more easily than T1 does no matter whether word construction is considered or not.

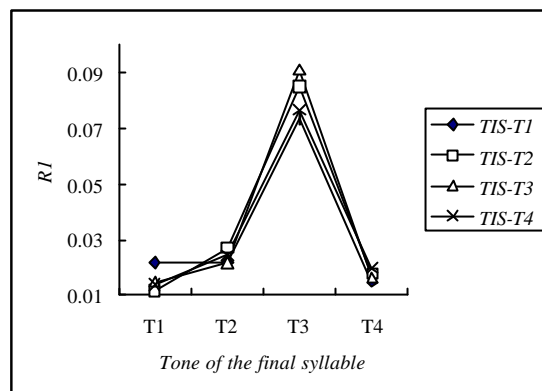


Figure 1(a): Tone pattern and the ratio of all the SU words

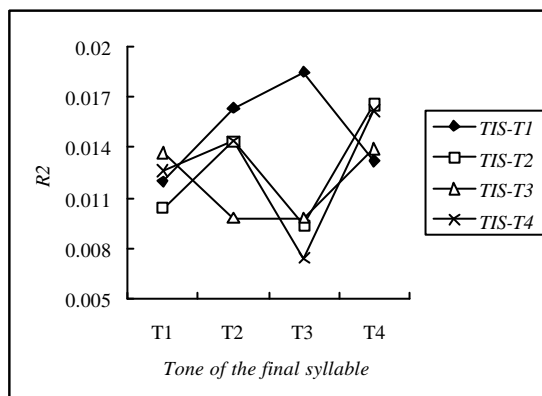


Figure 1(b): Tone pattern and the ratio of the SU words in which the unstressed syllables are roots

3. The relationship between tone pattern and word stress in continuous speech

In this section the results of our two previous studies with respect to word stress in continuous speech will be re-presented and re-analyzed. The materials and targets of these two studies as well as the present study are different from each other; however, the results of them are correlated.

3.1. Word stress perception in utterances

A perceptual experiment on word stress was conducted based on 1,898 Chinese disyllabic words in 300 utterances taken from the Chinese speech corpus of Microsoft Research Asia [12]. In the corpus, all of the combinations “T3+T3”

were labeled as “T2+T3” because of the tone sandhi rule that T3 metamorphoses into T2 before T3. 21 listening subjects were asked to judge that which syllable is relatively prominent within each of the words they heard, and then each syllable got a stress score that was equal to the amount of the subjects who judged this syllable was prominent. Obviously, the score reveals the perceptual stress degree of a syllable in a word. The detailed result can be seen from Fig. 2, which shows the stress scores of the final syllable under all the tonal conditions. The variance of the score of the initial syllable can be deducted directly from the variance of the score of the final syllable because the score of the initial syllable is equal to the difference between the number of the subjects and the score of the final syllable, and the score of the final syllable is related to the present study more closely than the initial one is, so we exclude the score of the initial syllable in the following analysis.

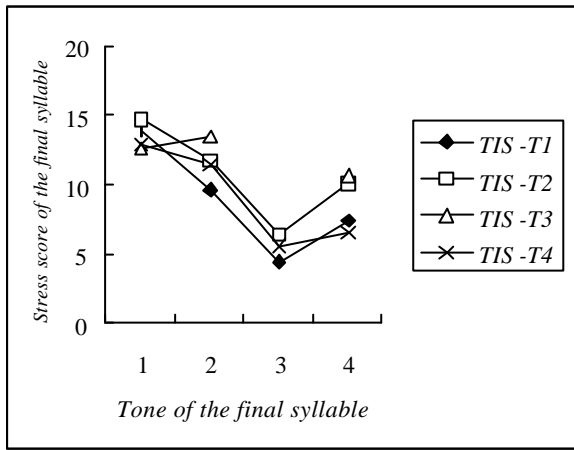


Figure 2: Tone pattern and perceptual stress score for the final syllable

As what we did in observing Fig.1, we also analyze Fig. 2 from both the final tone and the initial tone. Standing at the final tone, we can find out very easily that the final syllables bearing T1 got the largest stress scores, those bearing T2 got the second largest, those bearing T3 got the smallest scores, and those bearing T4 got the second smallest. That is to say, the final syllables with T3 are least prominent in perception, and those with T1 are most. The relationship among the stress scores under different initial-tone conditions is not as clear as it is under different final-tone conditions. As a whole, the final syllables under the initial-tone condition of T1 or T4 got a bit smaller stress scores than those under the initial-tone condition of T2 or T3 did. This implies that a syllable with T1 or T4 makes its following syllable less prominent than a syllable with T2 or T3 does. Yet the difference between the two initial-tone groups is quite small.

3.2. Sentence stress allocation within words

The words in the above experiment came from utterances, but that if a word obtained a sentence stress was out of consideration. In our subsequent study based on the same 300 utterances sentence stresses were labeled, being classified into two types, namely semantic stress and rhythmic stress. The reliabilities of both the classification and the labeling were proved by another perceptual experiment [13]. From the

results of the labeling, it was found that tone had an effect on the allocation of a sentence stress within a word, and the effect has nothing to do with the type of sentence stress[14]. Tab. 2 shows the detailed result, the percentages of the words in which the final syllables were allocated sentence stresses (final-stressed words) among the words with sentence stresses for each of the 15 tone combinations were listed, (a) being for semantic stress and (b) being for rhythmic stress. The percentages of the initial-stressed words are excluded in the following analysis for the reason similar to that why the stress scores of the initial syllables were excluded in section 3.1.

Each of the percentages itself is meaningless because we found that semantic stresses tend to locate on the initial syllables within disyllabic words, and rhythmic stresses on the final syllables[14]. What we should probe here is the relationship among the percentages under different tonal conditions.

Table 2: The percentages (%) of final-stressed words among the words with sentence stress for 15 tonal combinations

(a) For final-semantic-stressed words

Tone of final syllable \ Tone of initial syllable		Tone of final syllable			
		T1	T2	T3	T4
T1	60	7	0	7	
T2	82	12	0	14	
T3	32	31		13	
T4	50	16	0	5	

(b) For final-rhythmic-stressed words

Tone of final syllable \ Tone of initial syllable		Tone of final syllable			
		T1	T2	T3	T4
T1	97	80	23	43	
T2	100	90	30	83	
T3	92	100		87	
T4	93	86	31	47	

Observing from the final tone (by volumes), we can see that in both (a) and (b) the percentages of final-stressed words under the final-tone condition of T1 are the largest ones, those under the condition of T3 are the smallest, and those under the condition of T2 or T4 are in between. That is to say, within a word that carries a sentence stress, a syllable bearing T1 in the final position has the largest possibility to obtain the sentence stress, while that bearing T3 has the smallest.

The relationship among the percentages of the final-stressed words under the different initial-tone conditions is not very clear. The percentages under the initial-tone condition of T1 or T4 seem to be smaller than the corresponding ones under the condition of T2 or T3, yet a few exceptions in the table indicate that the effect of the initial tone on the allocation of sentence stress is not very significant.

What is more important than the orderliness itself in Table 1 is that the orderliness is corresponding to that in Fig.2. The most significant correspondence is that the tone that makes a syllable relatively prominent in a word also make the syllable

has relatively large possibility to be allocated a sentence stress, and the tone that makes a syllable relatively weak does the reverse. Specifically, within disyllabic words, the final syllables with T1 are most prominent in perception and have the largest possibility to obtain sentence stresses; on the contrary, the final syllables with T3 are least prominent and have the smallest possibility to obtain sentence stresses. It is also a similarity between the result in Fig.2 and that in Tab.2 that the effect of the initial tone on both word stress perception and sentence stress allocation within words is less significant than the effect of the final tone.

4. General discussion and conclusions

4.1. General discussion

In section 3 we are concerned with the relationship between tone pattern and word stress perception and the relationship between tone pattern and the allocation of sentence stress within words, and these contents are closely related with what we are concerned in section 2.

There is no question about that the engendering of SU words is a result caused by many linguistic factors including lexical and phonetic ones. From the viewpoint of phonetics, perception and production depend on each other. Thus, one can hypothesize that the syllables that are perceptually prominent have a smaller possibility to become neutralized ones than those that are perceptually weak do. Since words are mainly used in continuous speech, one can also hypothesize that the syllables that have relatively few opportunities to be allocated sentence stresses are easier to be unstressed at lexical level than those that have many opportunities. The main results of section 3 are, (1) the final tone has a closer relationship to word stress perception as well as sentence stress allocation than the initial tone does; (2) for the final tone, the syllables bearing T1 are most prominent in perception and have the most opportunities to be allocated sentence stresses, and those bearing T3 do the reverse. According to the hypotheses and these two results, the final syllables carrying T3 should be easiest to be neutralized at lexical level, and those carrying T1 should be most difficult.

Now we go back to see the results in section 2 again. The variance of $R1$ along with the variance of tone pattern coincides with the above deduction quite well. Though the variance of $R2$ does not completely accord with the deduction, it coincides the hypotheses partially, namely, the syllables with T1 are less easy to be neutralized than those with T4. That is to say, the hypotheses are confirmed.

Provided it is a linguistic universal that pitch is the closest correlate to word stress perception, the results of this study is reasonable phonetically. In Mandarin, T1 has only high feature, while T3 has only low feature in most cases, consequently syllables with T1 are most prominent in perception, and those with T3 are opposite. Though the target of the pitch curve of T2 and the beginning of T4 are high, the relatively low beginning of T2 and the low target of T4 degrade the prominence of the syllables bearing these two tones, thus, they are less prominent than the syllables bearing T1.

4.2. Conclusions

By analyzing the distribution of the original tones of the unstressed syllables at lexical level, the perceptual relative stress scores of syllables within words, the allocation of

sentence stresses within words, it is argued that tone has an effect on the engendering of unstressed syllables, on the perception of word stress pattern in continuous speech, and on the allocation of a sentence stress within a word. The mechanisms of tone effect on the engendering, the perception and the allocation are similar, and this implies that the relatively prominent syllable within a word is relatively easy to obtain a sentence stress and chronically the prominence of a syllable in continuous speech makes it uneasy to turn to a neutral syllable at lexical level. The final tone within a disyllabic word has a more significant effect on word stress pattern than the initial tone does. For the final tone, T1 makes a syllable to have the smallest possibility to be neutralized at lexical level, to be most prominent in word stress perception, and to have the largest possibility to be allocated a sentence stress, T3 does the reverse, and T2 and T4 do in between.

Acknowledgments

Work funded by NOCFL grant HBK01-05/013.

5. References

- [1] Chao, Y., 1979. *A Grammar of Spoken Chinese* (translated by Lv S.). Beijing: Commercial Press. 23.
- [2] Chao, Y., 1932. A preliminary study of English intonation (with its American variants) and its Chinese equivalent. *Journal of the Institute of History and Language of Central Academy (Supplement)*, S105-156.
- [3] Fry, D.B., 1958. Experiments in the perception of stress. *Language and Speech* 1(2), 126-152.
- [4] Bolinger, D.L., 1958. A Theory of pitch accent in English. *Word* 14(2-3), 109-149.
- [5] Lehiste, I., 1970. *Suprasegmentals*. M. I. T. Press. 125-132.
- [6] Lin, T., 1983. A Preliminary experimental study on the neutral tone in Beijing Mandarin. *Journal of Linguistic Study* 10, 16-37.
- [7] Wang Y., 2004. The effects of pitch and duration on the perception of the neutral tone in Standard Chinese. *Acta Acustica* 29, to be published.
- [8] Group of dictionary compilation, Institute of linguistics, Chinese Academy of Social Sciences, 1982. *A Dictionary of Modern Chinese*. Beijing: Commercial Press.
- [9] Li, W., 1981. A Study of the neutral tone and the normal stress in Mandarin. *Chinese Language and Writing* 1, 35-40.
- [10] Chao Y., 1933. Tone and intonation in Chinese. *Journal of the Institute of History and Language of Central Academy* 4(3), 121-134.
- [11] Liu, Y., Liang N., Wang D., et al., 1990. *A Dictionary of the Frequencies of the General Words in Modern Chinese*. Beijing: Yuhang Press.
- [12] Wang, Y., Chu, M., He, L., and Feng, Y., 2003. Stress perception of Chinese disyllabic words in utterance. *Chinese Journal of Acoustics* 22(1), 86-96.
- [13] Wang, Y., Chu, M., He, L., 2003. An Experimental study on classification and distribution of sentence stress in Mandarin. *Acta Psychologica Sinica* 35(6), 734-742.
- [14] Wang Y., Chu M., He L., 2003. Location of sentence stress within disyllabic words in Mandarin. In *the Proceedings of the 15th International Congress of Phonetic Science*. Barcelona. 1827-1830.