The Pitch Movement of Word Stress in Chinese

WANG Bei  ZHENG Bo  * LU Shinan  # CAO Jianfen  YANG Yufang

The Institute of Psychology & *The Institute of Acoustics, Chinese Academy of Sciences
# The Institute of Linguistics, Chinese Academy of Social Sciences

ABSTRACT

The pitch movement of word stress in Chinese is studied in aspects of word stress perception and pronunciation. Three parts of work have been done, word stress perception experiment, acoustic analysis of falling tone and the matched questions and statements experiment. The results of word stress perception experiment show that high point of the pitch range is the main acoustical correlate to word stress perception among low point, pitch range and average pitch. The variation of the pitch is not only correlates to word stress but also to the location of the syllable in a sentence, that is the intonation of the sentence. The acoustic analysis of falling tone is further studied in part II. In the matched questions and statements experiment, the stress level and the location of the target syllable are systematically regulated. The results of these two parts show that the pitch movement of stressed word is on basis of top and bottom-line intonation model. The pitch movement of high point is significant and free. The pitch movement of low point is limited by the bottom-line of intonation. The movement of high point is greater than that of the low point, which makes the pitch range vary.

INTRODUCTION

Sentence focus plays a central role in verbal communication. Speakers typically convey sentence focus by varying intonational attributes. In many cases, a focused word is accompanied by an increase in the fundamental voice frequency and duration (Atkinson, 1973; Brown and McGlone, 1974; O' Shaughnessy, 1979).

As a tonal language, the pitch movement of stressed word in Chinese is complicated that it can’t be described as one line intonation model. Chao proposed that the range of pitch between different tones and within the limits of moving tones is a variable quantity. The modification of the range is somewhat as a graph drawn on an elastic band would be magnified when stretched (Chao, 1933). The pitch range of tones can be described as the top and bottom-line model. The shifting up of the top-line correlates to the stressed word while the declination of the bottom-line keeps constant to make the rhythm of the sentence, which results in the widening of the pitch range (Sheng, 1995). Another hypothesis holds that the shifting of the frequency range in local tone contours can be regarded as the transposition or change-key of the melody in music (Wu, 1996). That is to say, the pitch movement of word stress is realized by shifting up of the pitch with constant pitch range.

Chinese TTS with high intelligibility is realized by concatenation of non-uniform units (Lu, 2000). Word stress regulation is important to improve the naturalness of synthesized speech.

The pitch movement pattern of word stress is studied in three parts of work. Firstly, the main pitch clue of word stress is studied by word stress perception experiment. The results show that the high point of the pitch is the main acoustic correlate to word stress among low point, pitch range or average pitch. Secondly, the acoustic analysis of falling tone is further done, since the pitch movement of word is not only correlate to word stress but also to the location of word in sentences, that is the intonation of sentences (Stephen et. al. 1986). Finally, the matched questions and statements experiment is done using meticulously designed materials. Both of the results of part II and part III show that the top and bottom-line declination intonation model does exist. The pitch movement of word stress is on basis of the intonation model.

1. WORD STRESS PERCEPTION EXPERIMENT

To find out the pitch movement of word stress, the first thing we should do is to study the acoustic correlate of word stress. It is well known that pitch, duration, and intensity are all word stress correlates. However, the pitch correlate to word stress is lack studied in Chinese. The word stress perception experiment is to study which one is the main clue to word stress perception, high point, low point of the pitch or pitch range.

Method

Materials

The pitch of Chinese syllables can not be described simply by average pitch because of tones. The pitch range of four tones is introduced in Chinese (Chao, 1933; Sheng, 1985). The high point, low point and pitch range are three indexes of pitch range.

Compare with other tones, the difference between high point and low point of falling tone is nearly the pitch range mentioned above. Words with falling tone are selected as materials in the experiment and falling experiments.

To study the relationship between pitch movement and word stress perception, the paired words are same Chinese characters and with same duration. To find out which one is the main clue to word stress perception, high point, low point and pitch range, four groups of paired words are designed (shown in figure 1). 1. high point identical; 2. low point identical; 3. pitch range identical and 4. pitch contour higher with smaller pitch range. Group 4 is especially used to find out whether high
point or pitch range is main clue to word stress perception.

All words are selected from a speaker’s large database. The high point, low point, pitch range and duration are measured using Multi-Speech software. 9 pairs of words are selected in group 1, 17 pairs in group 2, 37 pairs in group 3 and 2 pairs in group 4, totally 75 pairs.

![Figure 1](image)

**Figure 1** A Stressed words; 1. High points identical; 2. Pitch range identical 3. Low points identical; 4. High point higher with smaller pitch range.

B Unstressed word.

**Procedural**

After hearing the paired words, the subjects are asked to make a judgment which one is more stressed. To let subjects form a constant standard to word stress perception, 5 minutes exercise is done first. To deduce the sequence effect, the paired words are play twice as AB and BA and all paired words are played randomly. To make sure the judgment is reliable, the paired words are allowed to play again. All procedure are realized by computer programs using Visual Basic.

**Subjects**

16 female and 16 male subjects are from the Capital Normal University, Beijing. They are born and grown up in Beijing with normal hearing.

**Results**

1. The probabilities of selecting A and B as stressed syllable are shown in table 1.

<table>
<thead>
<tr>
<th>Condition</th>
<th>A</th>
<th>B</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>0.65</td>
<td>0.35</td>
<td>10.27**</td>
</tr>
<tr>
<td>1</td>
<td>0.50</td>
<td>0.50</td>
<td>-0.075</td>
</tr>
<tr>
<td>2</td>
<td>0.68</td>
<td>0.32</td>
<td>9.601**</td>
</tr>
<tr>
<td>3</td>
<td>0.69</td>
<td>0.31</td>
<td>11.032**</td>
</tr>
<tr>
<td>4</td>
<td>0.65</td>
<td>0.35</td>
<td>7.628**</td>
</tr>
</tbody>
</table>

* represents P<.05, ** represents P<.01.

2. The correlation between high point, low point and pitch range and the frequency the syllable been perceived as stressed is shown in table 2.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>High-point</th>
<th>Low-point</th>
<th>Pitch-range</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.352**</td>
<td>0.278**</td>
<td>0.189**</td>
<td></td>
</tr>
</tbody>
</table>

**Discussion**

A is significantly more perceived as stressed than B in group 2, 3, and 4. While the high point is identical between A and B in group 1, there is no significant percept difference. It can be concluded that the high point is the main clue to word stress perception. The results in group 1 and 3 does not support the low point is the main clue to word stress perception, since the height of the low point does not lead to correspond percept result. As for pitch range, the percept results in group 1 and 4 hold that the word with greater pitch range does not be perceived as stressed accordingly. The average pitch of B is higher in group 1, but B is not perceived as stressed. All the results in table 1 support that the high point is the main clue to word stress perception. As for low point, pitch range and average pitch, there is one or more results show that the height of low point, high point or average pitch does not lead to correspond percept result.

Moreover, the correlation between the frequencies the words been perceived as stressed and high point, low point and pitch range show that the correlation between the frequency and high point is highest (r=0.352**). The correlation between the frequencies and low point, and pitch range is significant is partly because high point, low point and pitch range vary the same time.

We can draw the conclusion that the high point is the main clue to word stress perception among low point, pitch range and average pitch. It is also true that the variation of high point is accompanied by low point and pitch range.

The pitch movement of word stress will be further studied in the following experiments.

**II. ACOUSTIC ANALYSIS OF FALLING TONE**

The paired words selected in experiment I are with similar duration and varied high point, low point or pitch range. The words with varied pitch is not only correlate to word stress but also to the location of words in sentences, namely at initial, middle or final of sentences. The pitch of word inlays in the intonation model. Considering the intonation model, 524 words in Part I are further studied in the acoustic analysis of falling tone.

**Methods**

**Materials**

524 words with falling tone are selected from a speaker’s large database.

**Words Labeling**

The location of the syllables in the sentence as at initial, middle or final is labeled. To syllables corresponding to the same Chinese character, the average high point and the low point of the pitch are calculated. While the high point of the syllable is larger than the average, the syllable is labeled as stressed, otherwise unstressed.

**Results**

1. The pitch movement patterns of stressed and unstressed syllables at initial, middle and end of the sentences are shown in figure 2.
2. Whether the difference of the high point, low point and pitch range are significant among locations are shown in table 3
3. Further ANOVA is done to separate the influence of word stress from location. The result is shown in table 4.
4. The correlation test of high point, low point of the pitch,
Acoustic analysis of falling tone uses word stress labeling according to object acoustic parameters. It is still necessary to study more speakers’ database. The way to word stress labeling is objective but there is more direct and nature way to compare stressed with unstressed word.

III. MATCHED STATEMENTS AND QUESTIONS

EXPERIMENT

The procedures used in this experiment involved the oral reading of sentences in which the location of sentence focus and the location of /dao4/ are systematically varied. The focusing of a word in a sentence is accomplished in a relatively natural way by asking speakers a question that prompts an answer having emphasis on one of the lexical items. The recorded sentences we analyzed using digital techniques, and statistical tests were applied to determine the effects of different focus locations on the fundamental frequency of selected key word, /dao4/.

Methods

Sentences Materials

There are 7+3=21 basic sentences with /dao4/ at the initial, the middle and the final of the sentences. The average length of the sentence is 7 syllables. For each basic sentence, two priming sentences are composed. Half of these priming stimuli were constructed to elicit responses with stressed /dao4/, and half with unstressed /dao4/.

Subjects and Recording Procedures

The speakers are 4 young men and 7 young women who can speak Standard Chinese well. Each speaker is seated in front of a microphone and given a stack of file cards with a typed sentence. On each trial, after listening to a prerecorded priming stimulus presented via a loudspeaker from computer, the speaker then read the sentence aloud into the microphone with the appropriate emphasis. The speaker’s responses are recorded to computer. The sampling rate is 11.25KHz.

Results

1. The correlations between the location of syllables and pitch movement of stressed syllables are shown in figure 3. Further significant tests are shown in table 6, 7 and 8.

Figure 3 The high point (A), low point (B) and pitch range (C) of stressed and unstressed word in different locations.

Table 6. The MONOVA of the high point, low point of pitch, and pitch range among locations × stress.

<table>
<thead>
<tr>
<th></th>
<th>High-point</th>
<th>Low-point</th>
<th>Pitch-range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>45.93**</td>
<td>50.88**</td>
<td>22.236**</td>
</tr>
<tr>
<td>Stress</td>
<td>100.09**</td>
<td>1.996</td>
<td>196.917**</td>
</tr>
<tr>
<td>Location × Stress</td>
<td>7.367**</td>
<td>1.998</td>
<td>9.243**</td>
</tr>
</tbody>
</table>
Table 7. The ANOVA of high point, low point of pitch, and pitch range among locations

<table>
<thead>
<tr>
<th></th>
<th>High-point</th>
<th>Low-point</th>
<th>Pitch-range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstressed</td>
<td>51.073***</td>
<td>33.363***</td>
<td>4.114*</td>
</tr>
<tr>
<td>Stressed</td>
<td>8.949***</td>
<td>18.887***</td>
<td>5.734**</td>
</tr>
</tbody>
</table>

Table 8. The ANOVA of high point, low point of pitch, and pitch range between stressed and unstressed

<table>
<thead>
<tr>
<th></th>
<th>High-point</th>
<th>Low-point</th>
<th>Pitch-range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>12.262***</td>
<td>0.445</td>
<td>38.481***</td>
</tr>
<tr>
<td>Middle</td>
<td>23.324***</td>
<td>1.229</td>
<td>10.336**</td>
</tr>
<tr>
<td>Final</td>
<td>87.13***</td>
<td>6.136*</td>
<td>32.755***</td>
</tr>
</tbody>
</table>

2. The correlation test of high point, low point, pitch range and duration of all speakers is shown in table 9.

Table 9 The correlation test of acoustic parameters

<table>
<thead>
<tr>
<th></th>
<th>High point</th>
<th>Low point</th>
<th>Pitch range</th>
</tr>
</thead>
<tbody>
<tr>
<td>High point</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low point</td>
<td>0.752**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Pitch range</td>
<td>0.775**</td>
<td>0.162**</td>
<td>1</td>
</tr>
<tr>
<td>Duration</td>
<td>0.150**</td>
<td>-0.173**</td>
<td>0.257**</td>
</tr>
</tbody>
</table>

Discussion

The declination of high point and low point of pitch does exist so that the pitch movement of stressed syllables in on basis of the top and bottom-line intonation model.

To the stressed syllables, the high points of the pitch are significantly higher (F=100.09**) which is consistent with that in large corpus analysis. All 11 speakers articulate the stressed syllable in this way. The low points of pitch do not change much (F=1.996) to stressed syllables. It is inconsistent with that in large corpus analysis in which the low point is significant higher to stressed word (F=114.12**). That is to say, to stressed syllables, the pitch movement of low points is not as significant as that of high points.

There is significant difference of high points of the pitch of the stressed syllables among different locations (F=8.949**), which is inconsistent with that in large corpus analysis (F=1.407). There are 7 speakers whose pitch movement of high points of stressed syllables is significantly different among locations. Whether the difference of high points to stressed syllables among locations depends on speakers. The movement of high points of stressed syllables is not limited by the top-line of intonation.

When the syllables at the initial or middle of sentences, the pitch movement of low points to stressed syllables is not significant. When the syllables at the end of sentences, the low point is significantly higher to the stressed syllables. It is true to all 11 speakers and large corpus analysis. The difference of low points to stressed syllables among locations is significant (F=18.887**), which is consistent with large corpus analysis (F=7.875**). The movement of low point is limited by the declination of bottom-line of intonation.

The correlation test of acoustic parameters is almost the same as that in large corpus analysis except that the correlation between low point and pitch range is significant (r=0.163**), The correlation between high point and pitch is significantly high (r=0.752**). The result is consistent with that in word stress perception experiment that the high point is the main cue to word stress. While a word is stressed, the movement of pitch is mainly realized by the high point of the pitch, which results in the variation of pitch range.

CONCLUSIONS

The high point of the pitch is the main cue to word stress perception and pronunciation. The pitch movement of stressed word in on basis of top and bottom-line intonation model. The pitch movement of high point is significant and free. The pitch movement of low point is limited by the bottom-line of intonation. The movement of high point is greater than that of the low point, which makes the pitch range vary.

REFERENCES

2. Yuanren Chao, Tone and intonation in Chinese, 145th Meeting of the American Oriental Society, 1933
7. Zong-ji Wu: A new method of intonation analysis for Standard Chinese: frequency transposition processing of phrasal contours in a sentence, Analysis, Perception and Processing of Language, G. Fant et al. (Editors), Published by Elsevier Science B.V. All rights reserved, 1996
8. Shinan Lu, Yufang Yang, and Jianfen Cao: Prosodic control in Chinese TTS system, Proc. ICSLP2000, Beijing, October

ACKNOWLEDGEMENT

The authors should give great thanks to Ms. HE Lin, Mr. WANG Lei, professor GUO Chunyan, and professor GUO Dejun.

1 Further copy of the article; please write to Yang Yufang, the Institute of Psychology, Chinese Academy of Sciences, China, 100101 or send e-mail to yangyf@psych.ac.cn.