PITCH DECLINATION IN THE STATEMENT SENTENCES IN MANDARIN

WANG Anhong
Department of Chinese, Peking University, Beijing
anhongw@sohu.com

LU Shinan
Institute of Acoustics, Academia Sinica
Lu_shinan@21cn.com

Chen Ming
Beijing Infoquick Sinovoice Speech Technology Corp.
chenming@sinovoice.com.cn

ABSTRACT

In this paper, the pitch declination in Mandarin is further observed in different tones. The speech materials come from a large-scale speech database that is broadcasting style. The final conclusions are: 1. There is declination generally in Mandarin. 2. The bottom point of the dipping tone can be taken as the real bottom point of the tone range and as the standard to measure the bottom points of other tones.

1. INTRODUCTION

Pitch declination theory is one of the important phonetic theories in America. Meada (1976) and Pierrehumbert (1979) discussed about it and put forward the concept of top line, baseline and key point. Top line is the result of the connection of all peaks of the pitch contour, and the baseline the result of all valleys. Both lines tend to be decrease and the decreasing speed of the top line is greater than that of the baseline. At the coordinate with the X for time and the Y for pitch, the key point appears at the point of a quarter from the first peak to the last one and the pitch value is approximately equal to the median of them. Another theory is the breath-group theory. Lieberman proposes that the nature of intonation be represented in the breathing features and the basic unit of intonation is breath groups. At the end of each breath, the fundamental frequency and amplitude decrease sharply with the decreasing of the air pressure of the sub-glottis. Breath-group theory only observes the declination in the final of a sentence, but has a weak capability of predicting the intonation pattern.

The declination theory has a powerful explanation and it can predict the intonation pattern of the whole sentence. But because the final phonetic result is affected by phonological, syntactical and pragmatic requirement, and even by tones in Chinese, the ideal pitch declination cannot be observed on all occasions. Chiu-yu Tseng (1981) studied the Chinese intonation with the recorded 50 spontaneous speech and the same sentences in broadcasting style, and then observed the top lines, baselines and key points from the measured pitch contour. The result is: in the spontaneous speech corpus, only 20% conforming to the pitch declination theory, and the most (74%) conforming to the breath-group theory among which no regular pitch declination is observed in these sentences. It is demonstrated in the study that there is an obvious difference in speakers and speech styles. For example, in the broadcasting corpus of one of the speakers, about 75% conform to the pitch declination theory. Chiu-yu Tseng observed the interaction of tone and intonation, and pointed out that the peak value and the valley value of pitch contour is related to tones. Shen Jiong (1992), with the method of tone aggregation, connected the top line and baseline by the top point and the valley point of the tone range respectively, and thus gave a powerful explanation to the Chinese intonation phenomenon. Shen Jiong put forward that top lines and bottom lines have different functions: top lines mainly represent sentence stress and bottom line sentence rhythm. Wang Bei (2002) testified it with physiological and physical experimental methods and large-scale corpus and pointed out that the acoustic expression of Chinese stress is the rise of the top point and the prolonging of syllable duration and that it is not related with bottom point except in the dipping tone. The pitch declination and resetting are mainly represented in the bottom line of prosodic phrases and intonation phrases in Chinese.

2. EXPERIMENTAL OBJECTS AND METHODS


2.1 Experimental Objects

The experimental objects of the study are the sentences in the speech database, coming from the speech database of Beijing Infoquick Sinovoice. The corpus is based on the common sentences of People’s Daily and is recorded in the broadcasting style by a female speaker. In our study, 50 sentences are firstly selected from the database. No punctuation for pause appears in the sentence so that the speaker can separate long sentences naturally. By the consideration of maneuverability, we adopt prominent pause to segment intonation phrases, and 92 prosodic phrases are segmented from 50 sentences.

2.2 Experimental Methods

Firstly the pitch contour is extracted with the Praat phonetic analysis software. If the pitch value is not right, for example, where the octave and semi-octave value appear, it will be corrected by hand. On the basis of it, the bottom points of the foot of different positions are measured in 92 prosodic phrases. The positions are divided into three parts in a prosodic phrase: the starting position, the medium position and the ending position. That is to say, all the bottom points of the first foot in the phrase are taken as the data of the starting position, all the bottom points of the last foot as the data of the ending position, and the bottom points between them as the data of the medium position. Selecting the foot as the measuring units has two advantages: first there will be more data of the bottom points of certain position, and the increased amount of data can make the analysis more creditable. Second the declination in the foot is also observed.

The starting point of the rising tone, the bottom point of dipping tone and the ending point of falling tone contribute to the valley value of intonation contours. In this paper, we get the bottom points of different positions in different tones statistically, and put forward that the rising tone, the dipping tone and the falling tone have different bottom point values. Whether in the starting position, the medium position or the final position, the dipping tone has a lower bottom point than the rising tone and the falling tone. So we consider the bottom point of the dipping tone as the reference of the bottom point of the tone range. The bottom point of the rising tone in the starting and the medium position in the phrase is lower than that of the falling tone, and it is taken as the sub-bottom point of the tone range. The ending point of the falling tone changes a lot under the influence of contexts, and it can be made no use of in the determination of the bottom line except when it is in the final position of the sentence.

3. RESULTS AND DISCUSSION

3.1 The Declination Representation of the Bottom Points of Different Tones

Except the level tone that only carries the feature of high pitch, the average values of the pitch bottom points of the other three tones (the rising tone, the dipping tone and the falling tone) in three different positions are demonstrated in Figure 1.

![Comparison of bottom points of different tones in different positions](image)

Figure 1

Two conclusions can be drawn from the above figure: 1. Statistics shows that the bottom points of each of the three tones tend to decrease from the initial to the final in the phrase. 2. If the bottom points are taken as the measurement objects of the pitch, the intonations in Chinese statement, like those in other languages, have the tendency of pitch declination.

The rising tone, the dipping tone and the falling tone have different bottom point values. Whether in the starting position, the medium position or the final position, the dipping tone has a lower bottom point than the rising tone and the falling tone. So we consider the bottom point of the dipping tone as the reference of the bottom point of the tone range. The bottom point of the rising tone in the starting and the medium position in the phrase is lower than that of the falling tone, and it is taken as the sub-bottom point of the tone range. The ending point of the falling tone changes a lot under the influence of contexts, and it can be made no use of in the determination of the bottom line except when it is in the final position of the sentence.

3.2 The Determination of the Bottom Line in Natural Speech and the Observation Results

According to the statistical data of the bottom points of different tones such as the dipping tone, the rising tone and the falling tone above, the bottom line can be divided into two different types, that is, the bottom line as the result of the connection of two or more bottom points and that by the parallel movement of two or more sub-bottom points of prosodic phrases. And thus prosodic phrases are divided into four types: type A, the prosodic phrases with the bottom line in the state of declination; type B, the prosodic phrases with sub-bottom line in the state of declination; type C, the prosodic phrases with the bottom line not in the state of declination; type D, the prosodic phrases whose
bottom lines are hard to be determined. The analysis results of the 92 intonation phrases are shown in the following table.

<table>
<thead>
<tr>
<th>Type</th>
<th>Quantity</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>49</td>
<td>53.3</td>
</tr>
<tr>
<td>B</td>
<td>29</td>
<td>31.5</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>D</td>
<td>13</td>
<td>14.1</td>
</tr>
</tbody>
</table>

During the process of classification, two phenomena are observed: first, if the only two dipping tone or rising tone appear in the same foot in an intonation phrase, it is not ascribed to type A or B, and to D; second, if the rising tone appears in the middle position of a three-syllable foot and becomes the transition of the preceding and following syllable, it is not considered as a sub-bottom point.

The pitch figures of two types of intonation phrases A, B, having the declination, are given first, and then the sample of type C, the only type that doesn’t conform to the declination, is analyzed.

The sentence in figure 2, “分妈妈给我讲故事与我给妈妈讲故事两类”. It can be divided into two prosodic phrases according to the obvious pause in the sentence: “分妈妈给我讲故事” and “与我给妈妈讲故事两类”. Both of them are of A type. “给我讲” can be taken as a foot, in which “讲” is a dipping tone and “事” in “故事” is the falling tone in the ending position of the phrase. By connection of the bottom points of them, a bottom line is formed and shows the tendency of declination. In the latter sentence, “给” in “分妈妈给我讲故事” and “讲” in “与我给妈妈讲故事” are all dipping tone, and by connecting the bottom points of them, a bottom line is also formed and has a declination tendency.

In figure 4, there is an only sample that doesn’t conform to the pitch declination in the 92 intonation phrases and then it is ascribed to type C. Bottom points appear in the four syllables “ 举 ” “ 很 ” “ 法 ” “ 使 ” , among which the bottom point of “ 举 ” is lower than that of “ 使 ” . By listening to the speech record repeatedly, it can be felt that “ 举 ” is stressed. So it can be explained as a special representation of the bottom points of the dipping tone when they are stressed. If the bottom points of the stressed dipping tone are not taken as a typical one, the only counter-example does not hold true any more. That is to say, except the 13 D-type phrases that cannot get any of the bottom lines directly, the other 79 prosodic phrases all conform to declination of bottom lines.

Our experimental results in substance support the intonation theory of pitch declination and the pitch resetting.
We did a similar experiment to Chiu-yu Tseng (1981), but drew a different conclusion. The reasons are: 1. In our research, the bottom line is defined as the connection line of the bottoms points of each tone range, but not the connection line of the valley points of the pitch contour in natural speech. 2. The speech style is the broadcasting one, different from Chiu-yu Tseng (1981) which observed spontaneous speech. But in her comparative research, it can be seen that there are a lot of sentences in broadcasting speech conform to the pitch declination theory. 3. The variance of speakers, which is quite obvious. Although the objects of the two researches are of Mandarin, they are of different times and districts. We hope that further study will be done as to these aspects.

4. CONCLUSIONS

Two conclusions can be drawn from the results of the above natural speech analysis.

1. In the prosodic phrases in Chinese, there is the pitch declination of bottom points, and the result is the same as that of Wang Bei. Thus it is testified that Chinese has the universal features of pitch declination and resetting like other languages. But the declination in Chinese is mainly represented by bottom lines but not by top lines.

2. Under the influence of tones, the falling tone, the rising tone and the dipping tone have different level of valley points in the pitch contour. The bottom point of the dipping tone is the lowest, and it can be taken as the real bottom point of the tone range and as the standard to measure the bottom points of other tones. The bottom points of the rising tone is in the medium of that of the dipping tone and the falling tone and can be taken as the sub-bottom point. As to the bottom points of the falling tone, there are two points: its bottom points is high and it varies a lot according to the degree of stress. Because the falling tone is not stable, the bottom points of the falling tone cannot be taken as the important data to observe the pitch declination.

REFERENCES