A Seven-tone Dialect in Southern China with Falling-Rising-Falling Contour: A Linguistic Acoustic Analysis

Zhu Xiaonong, Zhang Caicai

Division of Humanities, the Hong Kong University of Science and Technology, Hong Kong

hmzxhu@ust.hk, hmcelia@ust.hk

Abstract

This paper aims to give a quantitative description of the linguistic acoustic properties, i.e., fundamental frequency and duration, of a seven-tone dialect in Qiyang County (QY hereafter), south of Hunan Province, China. There are seven surface contrasting tones in the QY dialect, a branch of Old Xiang group. No acoustic analysis has been done in the QY dialect so far. What this paper concentrates on is the falling-rising-falling (FRF hereafter) pitch in two of the seven tones, which has been reported nowhere else. Many parameters which have potential influence on the pitch contour are considered so as to specify the factors which contribute to the form of this FRF contour. Although emphasis is placed on the analysis of citation tones of monosyllables, evidence from disyllables will be provided as well to support the pitch contour found in monosyllables. But a further perception test should be conducted to decide if the FRF contour is linguistically significant.

Index Terms: seven tones, FRF, acoustic analysis, perturbations

1. Introduction

As with all such phonetic studies, phonological descriptions always come first (Ladefoged 1997: 138,139). The QY tonal system contrasts one level, one dipping, two FRF and three rising pitches. The only level pitched tone (Tone6), sounds to stand at the middle of the speaker's normal range, e.g. /pa33/ eight. The dipping tone (T1) starts at the middle range, goes down a little and then rises to the lower half of the upper range, e.g. /ti33/ a plant. Of the three rising tones, one (T2) has a pitch which onsets in the lowest part of the range and rises into the upper half of the lower range, e.g. /pa23/ to crawl; one (T7) starts at almost the same position like T2 but rises higher than T2, e.g. /ti24/ white; one (T3) lies at the highest position of all the tones and is shorter than any other tone, e.g. /pa45/ a handle. Of the two FRF tones, one (T4) starts at the middle, goes down into the upper half of the lower range, rises into the lower half of the upper range, and then falls back to the middle. The other (T5) starts at the lower half of the lower range, goes down into the lowest point, rises back to the upper half of the upper range and finally falls back to the middle, i.e. /pa2143/ to stop.

2. Procedure

The corpus was composed of unstopped syllables, which were combinations of either of the three plosives /p, t, k/ and either of the three vowels /a, i, u/ in seven tones. Two sets of such combinations are demonstrated in the table. A complete list of corpus was not provided here due to the limited space.

Table 1: A Sample of the corpus for elicitation.

<table>
<thead>
<tr>
<th>T1</th>
<th></th>
<th>T1</th>
<th></th>
<th>T1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T2</td>
<td></td>
<td>T2</td>
<td></td>
<td>T2</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td></td>
<td>T3</td>
<td></td>
<td>T3</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td></td>
<td>T4</td>
<td></td>
<td>T4</td>
<td></td>
</tr>
<tr>
<td>T5</td>
<td></td>
<td>T5</td>
<td></td>
<td>T5</td>
<td></td>
</tr>
<tr>
<td>T6</td>
<td></td>
<td>T6</td>
<td></td>
<td>T6</td>
<td></td>
</tr>
<tr>
<td>T7</td>
<td></td>
<td>T7</td>
<td></td>
<td>T7</td>
<td></td>
</tr>
</tbody>
</table>

The corpus was recorded by twelve middle-aged native speakers of the QY dialect, six male and six female, which complies with the preferred minimum of half a dozen speakers of each sex for phonetic data analysis required by Ladefoged (2003, 14). Three female speakers were excluded from acoustic analysis for their irregular voice cords vibration. Thus totally nine speakers were calculated, still three better than the minimum of six people for quantified phonetic work mentioned by Ladefoged (2003, 14).

3. Normalisation

The speakers' mean F0 values were log z-score normalized, which involved calculating the log values of a speaker’s mean F0 at 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100 %, and then subtracting each log value from their mean log value (the mean log value was averaged with the two sampling point, 0% and 100% excluded) and dividing by their standard deviation log (again the onset and offset sampling points were excluded) (Zhu, 2005). The nine speakers’ Lz-score normalized values were shown graphically in figure 1. The mean Lz-score normalized F0 values were plotted on this figure as a function of their normalized duration.

Figure 1: Mean Lz-score normalized F0 curves for the QY citation tones.
As Rose pointed out (Phil Rose, 1988:56), the relationship between F0 and pitch is very complicated and the mono-dimensional assumption which equates the two exclusively is indefensible.

As stated in the first part, T2, T3, and T7 are all rising pitches in perception, T1 is a smoothly dipping pitch, and T6 is level. Thus after taking perception factors into account the pitch curves of QY tonal system can be expressed in figure 2.

4. Analysis in Linguistic Phonetics

However, another important question remains unsolved: how is the FRF curve formed, i.e., which factors contribute to the forming of such a strange shape? A possible solution is to list all the factors that may have an influence on the frequency and then exclude irrelevant factors one by one. Then, maybe the key is among the remainder.

4.1. Possible factors

Generally speaking, consonant types (only pre-vocalic consonants are considered here since the corpus is composed of open syllables only), intrinsic vowel height, phonation types and duration all have potential influence on the F0. We will discuss the four parameters respectively in the following.

4.1.1. Consonant types

Phonetic studies by Hombert 1975a, 1977b showed that a relative higher F0 of a following vowel caused by a voiceless consonant, and a relative lower F0 of a following vowel caused by a voiced consonant are salient in both tonal and non-tonal languages (J-M. Hombert, J. Ohala, and W. Ewan, 1979:40). Data presented by Glover 1970 on Tibeto-Burman languages indicate that breathy voiced consonants are stronger F0 depressors than voiced obstruents.

As stated above, in QY tonal system, T1, T3, T4, T6 always appear with voiceless obstruents, while T2, T5, and T7 co-exist with slack voiced obstruents. The significant distinctiveness between the voiceless and the slack voiced consonants is in the VOT values. The VOT of the voiceless group is c. 7 ms, while the slack voiced group is much longer, about 20 ms, and there is a leakage of small amounts of airstream between the release of closure and the onset of voicing.

In order to show the interaction between consonant types and F0, acoustic data was calculated from the syllable /ti/ in both T4 (帝) and T5 (地) pronounced by nine QY speakers. The time scale was set at about 100 ms after voice onset because possible perturbations caused by pre-vocalic consonants usually wouldn’t extend beyond 100 ms (Hombert, 1979: 40). F0 was sampled at 10ms, 20ms, 30ms, 40ms, 60ms, 80ms, and 100 ms. Results indicate that the duration of the perturbations caused by prevocalic consonants is even shorter in the QY dialect (about 30ms).

![Figure 2: Normalised pitch curves of QY citation tones after perception factors are considered.](image)

Even after the perturbations caused by pre-vocalic consonants vanished, the difference between T4 and T5 was still sufficient (c. 50 Hz) to be perceived by a normal speaker (above the perception threshold), which means that pre-vocalic consonants are not the deciding factor for the forming of pitch curve. Moreover, both T4 and T5 begin with a falling contour; if it was assumed that pre-vocalic consonant types cause the forming of FRF curve, the vowel after voiceless consonants should show a falling contour while the vowel after slack voiced consonants should have a rising contour. But the F0 curves with voiceless and slack voiced consonants are both falling. The contradiction between assumption (consonant types are the cause of FRF curve) and observation (the curves of F4 and F5 are both falling) is obvious. Thus, pre-vocalic consonant types can’t explain why the FRF curve exists.

4.1.2. Intrinsic vowel frequency

There is a connection between intrinsic vowel frequency and the relative height of the average fundamental frequency: other factors being constant, higher vowels have higher fundamental frequency (Lehiste, 1970:68). But the intrinsic F0 associated with different vowel qualities is manifested as higher or lower F0 level throughout the whole vowel range. Thus, the intrinsic vowel frequency can be excluded from the affecting factors because it generally raises or lowers the overall pitch levels but supposedly has nothing to do with the circumflex of the curves.

4.1.3. Phonation types

Hombert (1977, p. 29) notes that languages with more than six tones usually recruit an additional phonetic parameter like phonation type into their tone system.

There exists one phonation type in QY phonological system, the slack voice, which is realized phonetically as a different type of obstruents, slack voiced obstruents. Like breathy voice, slack voice leads to a lower frequency at the vowel onset. As we have discussed the consonant types in section 4.1.1, it won’t be discussed here.

4.1.4. Duration

No doubt change in duration will cause a change in F0 as well. But in most cases, duration is merely a concomitant parameter...
of F0; for instance, circumflex curves tend to have a longer duration while falling pitches may be potentially shorter.

No significant variation in duration is observed in QY tonal system. The average duration values for nine QY speakers indicate that circumflex curves (T1, T4, and T5) are longer than non-circumflex curves (including one level pitch T6, and three rising pitches T2, T3, T7). Yet it can’t be inferred from this observation that longer duration makes the pitch curve become FRF although a longer duration may foster a more circumflex pitch curve through providing more space for a pitch to maximally realize its circumflex trajectory.

4.2. Evidence from tone sandhi in disyllables

The spectrograms of some disyllables indicate that, no matter whether the FRF curve appears in the first or the second syllable, the FRF curve is always apparent. Figure 4 shows the waveforms and spectrograms of two disyllables with T4 at either the former or the latter position. Here the two windows both stretch for 1s with the blue lines indicating the F0 curve.

![Figure 4: Upper figure: /g289/g10450/g11803 [t/g1190 Æ3243 p/g1190 i/g446 21] by M4 (TZL). Lower figure: /g18080/g8073 [g1102 ia/g446 334 g1102 in3243] by the same speaker.]

4.3. Summary

According to the above analysis, no factor can sufficiently explain the forming of the FRF curve. Of all these parameters, pre-vocalic consonant types and phonation types refer to the same thing, but their strength over F0 don’t extend beyond 100 ms after the vowel onset; intrinsic vowel frequency and duration are concomitant factors which can’t modulate the circumflex curve of F0. Evidence from tone sandhi contour also indicates that the two pitches (T4 and T5) are in themselves FRF curves.

Yet FRF phenomenon is so complicated that many questions still remain to be discussed, like the influence of intonation on pitch curve. Since all the citation forms were recorded separately, each syllable may have carried the effect of intonation, i.e., the falling ending is caused by the famous phenomenon of declination. However, the current evidences tend to support the existence of the FRF curve.

5. Further work

In order to ascertain if the falling curve at the end of the range is linguistically significant, frame sentence can be used to eliminate the effect of intonation.

6. Acknowledgements

This paper owes thanks to the following two grants: Hong Kong RGC CERG Grant (HKUST 6426/06H), and HKUST Research Project Competition Grant (RPC06/07.HS03).

7. References