A Comparative Study between Intonation Question and Particle Question in Cantonese on Their Realization of $F_0$ Contours

Wentao Gu, Keikichi Hirose and Hiroya Fujisaki

University of Tokyo, Japan

{wtgu; hirose}@gavo.t.u-tokyo.ac.jp  fujisaki@alum.mit.edu

Abstract

Question has a significant effect on $F_0$ contours in various languages, among which tone languages require more careful study because of the complex interaction between lexical tones and phrase intonation. In this study we employ the command-response model to compare the realization of $F_0$ contours in two types of questions in Cantonese, a typical tone language with nine lexical tones. In intonation question, the tone command in the later part of the sentence-final syllable is substituted by a positive tone command. Particle question, on the other hand, maintains the inherent tone command pattern for the sentence-final particle. In both cases, the effect of question on $F_0$ contours is localized in the sentence-final part, not only by an addition of sentence-final phrase command, but also by the question ending tone command, the absolute amplitude of which is usually larger and also indicates the degree of inquisitive intention.

1. Introduction

An accurate and quantitative representation of the essential characteristics of $F_0$ contours of speech is necessary for synthesis of expressive speech. The $F_0$ contour of speech is affected by complex combination of various linguistic, paralinguistic and non-linguistic factors [1]. Among them, speech-act type, as a quite fundamental factor in information transmission, is well known to have significant effect on $F_0$ contours of speech, though the effect may vary with languages.

The best-known contrast in speech act is between statement (declaration) and yes/no question (interrogation), which in many languages are associated with sentence-final $F_0$ falling and rising respectively. Many works have been done on questions in various languages, among which tone languages require more careful study because of the complex interaction between phrase intonation and lexical tones which have an important role to distinguish word meanings. In this study, we investigate specifically the effect of questions in Cantonese, a typical tone language well-known for its complex tone system.

Cantonese has many types of yes/no questions [2], but we only study the following two typical categories in this paper:

1) The ‘intonation question,’ which is identical to a statement in the text, but differs from a statement only in intonation. For example, the so-called echo question, which is a repetition of a statement out of surprise or suspicion, is a type of intonation question.

2) The ‘particle question,’ which is identical to a statement in the text except for an addition of a specific question particle at the end. Cantonese is especially rich in particles [2, 3], and unlike Mandarin where particles are usually of neutral tone, particles in Cantonese can carry various lexical tones.

Intonation question is paralinguistic, whereas particle question lies between linguistic and paralinguistic. Previous studies on Cantonese questions have shown that the intonation question modifies the $F_0$ curve in the sentence-final syllable into a rising shape regardless of the tone identity of the sentence-final syllable [3, 4]. However, these observations and analyses were phenomenological and lacked the insights into the underlying mechanisms, because they were not based on quantitative models for the process of $F_0$ contour generation.

By introducing the command-response model [1] for Cantonese [5, 6], we can use a different approach to study the effect of questions on $F_0$ contours of Cantonese. Our recent study [7] has shown that the intonation question substitutes the tone command in the later part of the sentence-final syllable by a positive command with larger amplitude. This paper will continue to make a deeper discussion and give a comparative study between intonation question and particle question.

2. The command-response model for $F_0$ contour generation for Cantonese

The command-response model for tone languages describes $F_0$ contours in the logarithmic scale as the sum of phrase components, tone components and a baseline level $\ln F_0$. The phrase commands produce phrase components through the phrase control mechanism, giving the wide-range shape of $F_0$ contour, while the tone commands of both positive and negative polarities generate tone components through the tone control mechanism, characterizing the local $F_0$ changes. Both mechanisms are assumed to be critically-damped second-order linear systems. For a specific tone language, a set of tone command patterns needs to be specified in the model.

Cantonese has a complex system of nine lexical tones, as shown in Table 1. The syllables of entering tones end with an unreleased stop coda /p/, /t/, or /k/, and hence are shorter than their counterparts of non-entering tones. We have proposed the tone command patterns for the nine tones in Cantonese [5, 6]. In terms of the command polarities corresponding respectively to the earlier and the later parts of a syllable, the command patterns for each tone can be represented phonologically as in the rightmost column of Table 1, where ‘+’ and ‘0’ denote positive and null commands respectively, while ‘-’ and ‘=’ both denote negative commands but ‘=’ stands for a more negative one than ‘-’ and the brackets indicate entering tones.

<table>
<thead>
<tr>
<th>Tone name in Middle Chinese system</th>
<th>Tone number</th>
<th>Pitch feature</th>
<th>Tone code</th>
<th>Command pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-entering tones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper-level</td>
<td>T1</td>
<td>high level</td>
<td>55</td>
<td>++</td>
</tr>
<tr>
<td>Upper-elevating</td>
<td>T2</td>
<td>high rising</td>
<td>35</td>
<td>-</td>
</tr>
<tr>
<td>Upper-departing</td>
<td>T3</td>
<td>mid level</td>
<td>33</td>
<td>0 0</td>
</tr>
<tr>
<td>Lower-level</td>
<td>T4</td>
<td>low falling</td>
<td>21</td>
<td>= =</td>
</tr>
<tr>
<td>Lower-elevating</td>
<td>T5</td>
<td>low rising</td>
<td>13</td>
<td>0 0</td>
</tr>
<tr>
<td>Lower-departing</td>
<td>T6</td>
<td>low level</td>
<td>22</td>
<td>- -</td>
</tr>
<tr>
<td>Entering tones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper-entering</td>
<td>T7</td>
<td>high level</td>
<td>5</td>
<td>[+ +]</td>
</tr>
<tr>
<td>Middle-entering</td>
<td>T8</td>
<td>mid level</td>
<td>3</td>
<td>[0 0]</td>
</tr>
<tr>
<td>Lower-entering</td>
<td>T9</td>
<td>low level</td>
<td>2</td>
<td>[- -]</td>
</tr>
</tbody>
</table>
Among the nine tones, T3 and T8 have no tone commands, and T2 has a pair of tone commands, while all the other tones actually possess a single tone command. It is to be noted that T4 and T6 show the same command polarities but T4 gives more negative amplitude than T6 [5, 6]. The command patterns of entering tones are similar to those of their respective counterparts of non-entering tones, except for a shorter command duration caused by the unreleased stop coda.

3. Speech data

Two sets of speech materials are used: Speech Material I for intonation question, while Speech Material II for particle question. Each material consists of two subsets: A for studying qualitative distinction between question and statement, while B for studying quantitative distinction in question.

Speech Material I-A consists of two groups of sentences, each of which shares a respective fixed carrier frame as below:

"Nei1 go3 zi6 duk9 " (This character is read as __)
"Sin1 duk9 __ nei1 go3 zi6" (First read the character __)

In both of them, a target syllable /maa/ or /ma(a)k/ carrying each of the nine tones, is embedded at the underlined position. Each sentence is uttered in two different modes: neutral statement and intonation question, as shown below.

- I-A-1: “Nei1 go3 zi6 duk9 __”
  “Nei1 go3 zi6 duk9 __”
- I-A-2: “Sin1 duk9 __ nei1 go3 zi6.”
  “Sin1 duk9 __ nei1 go3 zi6.”

From material I-A-1 the interaction between question intonation and sentence-final tones can be investigated. However, considering that the embedded syllables are liable to unconscious emphasis and hence the results on material I-A-1 may be confounded with the effect of emphasis, material I-A-2 is also introduced to confirm the results by placing the embedded syllables apart from the sentence-final position.

The intonation question can also be quantitative, viz., it can be uttered with different degrees of doubt on the subject. To investigate the effect of such quantitative distinction, three versions of question utterances using the same carrier frame as in material I-A-1 but only with the embedded syllable /maa3/) expressing progressively increasing degrees of doubt are collected in Speech Material I-B.

Speech Material II-A consists of the following three carrier frames ending with different question particles:

- “Nei1 go3 zi6 duk9 __ me1?” (surprise or disbelief)
- “Nei1 go3 zi6 duk9 __ gwaas3?” (expressing uncertainty)
- “Nei1 go3 zi6 duk9 __ aa1?” (asking for confirmation)

The three question particles in different tones express the intentions as indicated in the parentheses respectively. In fact, the three questions indicate negative, neutral, and positive presuppositions respectively, showing the progressively increasing confidence in the subject. They correspond approximately to the intonation questions with progressively decreasing degrees of doubt. The target syllables embedded in these sentences are the same as in Speech Material I-A.

Among these three particle questions, the question ending with the particle /me1/ can further be quantitatively distinguished by different degrees of negative presupposition. Therefore, three versions of question utterances (only with the embedded syllable /maa3/) expressing progressively increasing degrees of negative presupposition are collected in Speech Material II-B to investigate this quantitative distinction.

The informant is a male native speaker of Cantonese from Guangzhou. Each sentence was uttered four times at the normal speech rate of the speaker. Therefore, there are altogether 72, 72, 12, 108, and 12 utterances in Speech Materials I-A-1, I-A-2, I-B, II-A, and II-B respectively.

The F0 contours of these utterances were analyzed through Analysis-by-Synthesis, viz. the parameters were first estimated manually by aid of linguistic information and then optimized by a hill-climbing procedure of successive approximation.

4. Intonation question

The data in Speech Material I-A-1 show that intonation question results in an F0 rising in the sentence-final part regardless of the sentence-final tone identity. The detailed results of Analysis-by-Synthesis have been given in [7] and will not be repeated here. Figures 1 and 2 show the results on F0 contours of a set of utterances: Fig. 1 for neutral statement and Fig. 2 for intonation question. Due to space limitations, here only the utterances with embedded T3, T4 and T6 target syllables are shown in panels (a) ~ (c) respectively.

The downdrift of F0 in declarative utterances of Cantonese as reported in previous works [2] is a natural result of the phrase control mechanism in the framework of the command-response model. This is clearly shown from the examples.

Comparison between the two figures indicates that the difference between F0 contours of statement and intonation question is highly localized in the sentence-final part. The analysis shows that in intonation question a positive tone command needs to be introduced in the later part of the sentence-final syllable (henceforth we call it question ending tone command) regardless of the tone identity of this syllable. The tone command pattern in the earlier part of the same syllable, however, remains unchanged for any tone identity.

A detailed discussion in our recent work [7] has revealed that the question ending tone command acts as a substitution (instead of an addition) of the inherent tone command, with a certain degree of compensation effect: the lower the preceding

---

**Figure 1:** Analysis-by-Synthesis of F0 contours of the utterances of statement in Speech Material I-A-1.
tonic command in the early part of the sentence-final syllable is, the higher the question ending tonic command tends to be.

Meanwhile, compared with the statement, an additional phrase command is also observed before the end of the intonation question. However, we cannot claim that this is an effect of intonation question. There may be an interaction with the effect of emphasis [7], because the sentence-final syllable here is an embedded target which may receive more emphasis unconsciously in an intonation question than in a statement.

In order to clarify this problem, we also investigate Speech Material I-A-2, in which the target syllables are embedded at a sentence-medial position. Analysis shows a similar result as in I-A-1, indicating that addition of a sentence-final phrase command is an essential feature of intonation question. Figure 3 gives the analysis results on the $F_0$ contours of two example utterances in which the syllable /maa3/ is embedded, panel (a) for statement and panel (b) for intonation question. It shows clearly that the phrase command immediately before the sentence-final compound word “nei1 go3 zi3” (this character) is larger in intonation question (0.15) than in statement (0.07).

Finally, Speech Material I-B is analyzed to investigate the quantitative distinction in intonation question. Figure 4 gives the results of Analysis-by-Synthesis on $F_0$ contours of three utterances of intonation question which share the same embedded syllable /maa3/ but express progressively increasing degrees of doubt. Analysis shows that the higher the degree of doubt is, the higher rise of $F_0$ tends to occur in the sentence-final part, and hence both the larger question ending tone command (0.43, 0.67, and 0.86 in panels (a) ~ (c) respectively) and the larger sentence-final phrase command (0.13, 0.21, and 0.40 in panels (a) ~ (c) respectively) tend to be given.

5. Particle question

Analysis of the data in Speech Material II-A shows that the difference between $F_0$ contours of particle question and statement is also highly localized in the sentence-final part – especially within the question particle, of which the inherent tone command pattern is maintained. This is clearly shown in Fig. 5, where the $F_0$ contours of three question utterances sharing the same embedded syllable /maa3/ but ending with particles of different tones are analyzed. The $F_0$ contours of these three utterances all show a sentence-final phrase command which is absent in the statement, but they differ in the tone command corresponding to the sentence-final particle.

Unlike in intonation question, the sentence-final tone identity in particle question is maintained, but the absolute amplitude of the tone command (if it exists) in the question particle is much larger than usual. When there is no tone command in the question particle (e.g., /gwaa3/), the syllable duration of the particle tends to be longer than usual. This can
6. Discussion and conclusion

By use of the command-response model, we have compared the realization of $F_0$ contours between intonation question and particle question in Cantonese. Both types of questions add or enhance a phrase command in the sentence-final part, but they differ in the pattern of the ending tone command. Intonation question results in an $F_0$ rising at the sentence-final position, regardless of the ending tone identity. Although previous works [3] showed the same observations, they did not provide a fully quantitative analysis by separating local tones from phrase intonation. On the basis of autosegmental phonology, Law [3] claimed that in intonation question the tones in the sentence-final syllable are converted either into high level tone (T1/T7) or into high rising tone (T2). However, her analysis is too simple to be accurate for the following two reasons. First, it ignores the difference between the patterns ‘0’, ‘0+’, ‘-’ and ‘+’ (they are all regarded as T2 there). It can be clearly seen from Fig. 2 that the sentence-final syllables of T3, T4, and T6 in intonation question show the three different patterns respectively. Second, it also ignores the difference in command amplitude. For a sentence-final syllable of T1 in the intonation question, the tone command in the later part of the syllable is much larger than that in its earlier part, and hence the tone pattern here differs from the inherent pattern of T1.

Our analysis shows that the inherent tone command in the earlier part of the sentence-final syllable is maintained, whereas the one in the later part is substituted by a positive tone command with larger amplitude than usual. Also, its amplitude is approximately in reverse proportion to that of its preceding tone command so as to make certain compensation.

Particle question, on the other hand, keeps the inherent tone command pattern for the sentence-final particle, except that the question particle has a tone command of larger absolute amplitude or a longer syllable duration than usual.

Moreover, both types of questions can convey different degrees of inquisitive intention. Intonation question realizes it by different amplitudes of question ending tone command, whereas particle question realizes it not only linguistically by different question particles (hence different polarities of the sentence-final tone command) but also paralinguistically by different amplitudes of the tone command in the particle.

7. References