ISCA Archive
http://www.isca-speech.org/archive

TAL 2006 – International Symposium
on Tonal Aspects of Languages
La Rochelle, France
April 27-29, 2006

A Phonetic Study on the Acquisition of Cantonese Tone

Ka Yan Margaret LEI

Department of Linguistics and Modern Languages
Chinese University of Hong Kong
margaret@cuhk.edu.hk

Abstract

This study adopts an acoustical approach in analyzing the fundamental frequency (F_0) patterns in tonal production of a Cantonese-speaking child, based on audio data from the Hong Kong Cantonese Child Language Corpus (CANCORP) [5]. Through examining F_0 patterns and variability on the production of different types of tones in Cantonese, the study attempts to address issues in the acquisition of tone by Cantonese-acquiring children.

1. Introduction

1.1. Tone system in Cantonese

Most scholars considered Cantonese as having six distinct tones [1, 2, 4]. There are three level tones: T1 (high level, 55), T3 (mid level, 33), and T6 (mid-low level, 22), two rising tones: T2 (high rising, 25) and T5 (mid-low rising, 23), and one falling tone: T4 (mid-low falling, 21). There are also three entering tones which correspond in pitch height to the three level tones, high (T7, 5), mid (T8, 3) and mid-low (T9, 2), which consist of syllables ending in –p, -t, or –k respectively.

1.2. Acquisition of tone

Not many studies on the phonological acquisition in Cantonese, especially on tones, have been carried out as compared to other aspects of language development. Earlier studies focused on Chinese, with the first of the kind by Chao [3] on his Mandarin-speaking granddaughter at the age of twenty-eight months. It was observed that tones were acquired very early. Longitudinal studies in Mandarin [7] and Cantonese [8] have also found that the period of tone acquisition was relatively short. Later studies on a variety of tonal languages like Mandarin [7], Cantonese [8], Thai [10], and Lao [11] further confirmed that correct tone production was acquired well before segmental production.

1.3. Tone production in Cantonese-speaking children

According to Tse [8] and Tse [9], the high level tone (T1, 55) was the earliest appeared tone in Cantonese-acquiring children, to be followed by the mid level tone (T3, 33) and the high-rising tone (T2, 25). Despite of some minor discrepancies on the actual order of appearance of the other tones between the two studies, the three entering tones agreed to appear in the same chronological order from high (T7, 5), mid (T8, 3) to mid-low (T9, 2).

The phonetic study on tones by Lee [6] showed that the F_0 contours of the nine citation tones in children aged 9- to 10-years old were similar to those of adults, except that the absolute frequencies among the speakers differ.

1.4. The current study

Despite the fact that a number of studies on the acquisition of tone have demonstrated an early mastery of tone in children acquiring a tonal language, most of them adopted an impressionistic judgment on the accuracy of tone production. As a suprasegmental feature, it is unavoidable that lexical information could possibly interfere with the interpretation of the tonal value of a syllable, meaning that it is possible for investigators to be biased in their judgment and bounded to map the tone value of a syllable with the readily available ones in the lexicon. In addition, the rich inventory of tones in Cantonese could have further complicated the issue, as the three level tones differ almost minimally in their absolute pitch heights only and the two rising tones have almost identical F_0 onset and are more distinguishable at the F_0 offsets. Thus, these all put doubt to the claim on early mastery of tone production.

On the other hand, most of the acoustical research in speech development have focused on children above three years old, leaving the picture of children below three years old less explored, especially on children acquiring a tonal language.

The present study aims to provide a preliminary account on the acquisition of tone in a child acquiring Hong Kong Cantonese as his native language at the age of two years seven months old with the aid of acoustical tools for analysis. We attempt to examine patterns of F_0 distribution and pitch contours among the different Cantonese tones, and the corresponding degree of variability.

2. Method

2.1. Corpus

Speech data was taken from the Hong Kong Cantonese Child Language Corpus (CANCORP) [5], which provides longitudinal records on the early language development of Cantonese-speaking children.

2.2. Speech samples

Acoustical analysis was made based on the speech samples from the original audio recordings of a child (CKT) acquiring Hong Kong Cantonese as his native language, at the age of two years seven months. The recordings were conducted at regular intervals under naturalistic settings with the parent and the investigator interacting with the child.

2.3. Data selection

All utterances produced by the child, including both spontaneous and imitated productions, which are audible, recognizable and non-overlapping with other sound sources were first segmented into separate digitized audio files.

The process of fine data selection was then carried out through careful examination of the segmented utterances, based on several major criteria in discarding unsuitable syllables. Several types of syllable were not included in the analysis: (i) unclear words: whose phonetic properties largely differ from those available in Cantonese lexicon and their lexical status could hardly be determined from possible contextual environment; (ii) words or utterances in strong emotion: as indicated by exaggerated duration and/or pitch contour; (iii) sentence-final particles: which incorporate strong effect of intonation. In cases when there coexisted any audible extra sound source in addition to the normal background noise with some parts of the utterances of the child, the affected syllables were discarded from analysis to avoid possible effects of the undesired sound source on the phonetic analysis of the targeted syllables. Common extra sound sources included the preceding or following utterances of the investigator or the parent, or the sounds from toy playing.

2.4. Syllable segmentation and data processing

Syllable segmentation and data processing were carried out in PRAAT. Before the labeling of syllables and marking of syllabic boundaries, adjustment was made by editing the missing or incorrectly marked cycles of glottal pulses manually before data processing. Each segmented syllable was time normalized and divided into twenty equal portions. The average F_0 value of each portion was obtained with PRAAT.

It is observed that the waveforms of speech produced by children are generally less regular as compared with the ones produced by adults. The degree of irregularity and fluctuation gradually diminishes with age. In our study, syllables were discarded from analysis under conditions when their corresponding waveforms appeared to be too irregular and could not provide indicative cues about their cycles, thus affected the accuracy of glottal pulse marking by PRAAT. Due to considerable amount of missing or incorrectly marked cycles, the resultant F_0 were always of extreme values. These happened in situations when the child spoke with a cracked voice or in a shouting manner.

2.5. Data analysis

Pitch contour of each syllable was analyzed with the aid of two types of pitch curves plotted. The first type of curve was plotted with reference to the average F_0 values from the twenty sample portions. Tone normalization was carried out with the logarithmic z-score (LZ) transforms [13]. The resultant normalized values in terms of logarithmic z-scores were used to plot the second type of curve. Formula (1) was used to compute the normalized logarithmic z-score value from the original F_0 values:

$$z_i' = \frac{\log_{10} x_i - \frac{1}{n}\sum_{i=1}^{n}\log_{10} x_i}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\log_{10} x_i - \frac{1}{n}\sum_{i=1}^{n}\log_{10} x_i\right)^2}} \quad (1)$$

The LZ transformation expresses F_0 value as a multiple of measure of dispersion away from a mean value, all of which are in logarithmic terms [13].

3. Results and Discussion

3.1. Overall fundamental frequency distribution

The overall mean F_0 from the average pitch values of all the sample portions was 327.82Hz, having 235Hz as the minimum F_0 and 511.10Hz as the maximum F_0. The standard deviation of these portions was 55.16Hz.

3.2. Descriptive analysis on individual types of syllables

3.2.1 F_0 variability

For the same type of CV syllables, it is observed that the F_0 contours varied noticeably in different utterances, even when the neighboring contextual environment remained almost the same. An example is depicted in Figure 1. In this example, the first syllable *ngo5* is in the sentence-initial position and *co5* is the following syllable of it. As we can see from the three curves, the F_0 contours differed substantially in terms of both contour shape and the relative pitch height, especially for the second syllable *co5*.

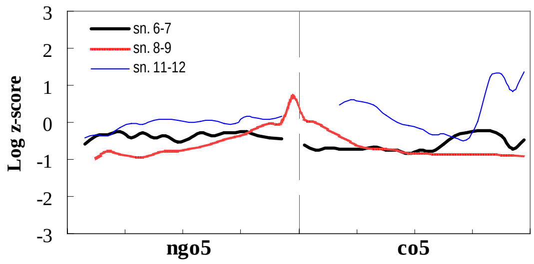

Figure 1. *F_0 curves of 3 instances of "ngo5 co5"*

For the ease of data processing, each segmented syllable was labeled with a Syllable Number (sn.) in chronological order of occurrence.

Another example is depicted in Figure 2 where the absolute pitch height of the three high level tone words *hoi1* varied from each other with more than one unit of log z-score.

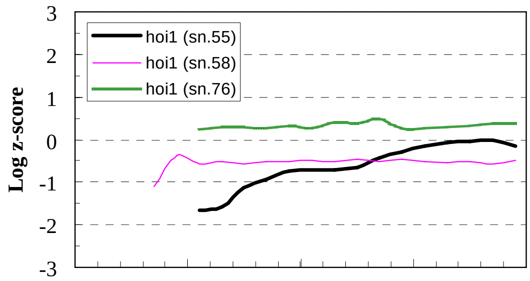

Figure 2. *F_0 curves of 3 "hoi1"*

In addition to the large variability in F_0, the F_0 curves are also observed to be more fluctuating as compared with adults' production.

3.2.2 *The 3 level tones*

Among the three level tones, F_0 values for most of them clustered around the central portion within the pitch range of the child. There is no clear separation for distinguishing among the three level tones. Figure 3 shows the F_0 contours of six syllables, with two syllables of the high level tone, the mid level tone, and the mid-low level tone respectively. It can be observed that the F_0 curves of the three types of level tones overlap with each other to a considerably great extent.

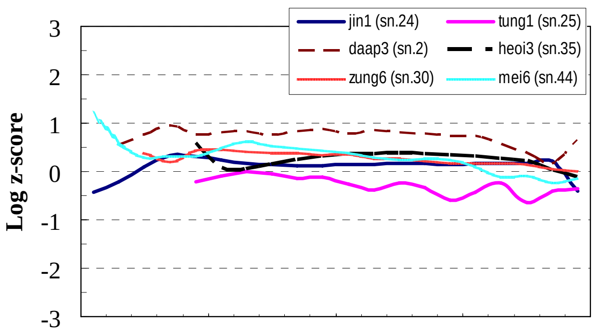

Figure 3. *F_0 curves of the 3 level tones*

3.2.3 *The 2 rising tones*

For the two rising tones, the rising property is not clearly demonstrated in the F_0 contours of the corresponding syllables. As shown in Figure 4, only one instance of *ngo5* has a noticeable rising contour. There is also no clear distinction in the F_0 contours between the high-rising and the low-rising tones.

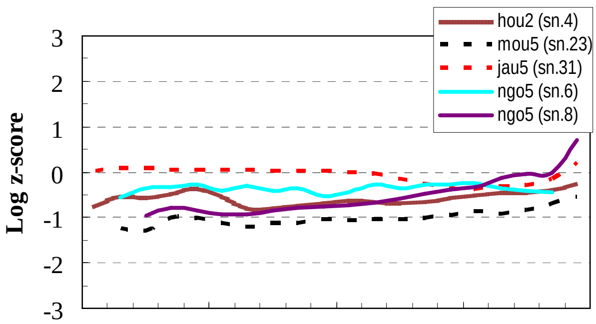

Figure 4. *F_0 curves of the 2 rising tones*

3.2.4 *The falling tone*

Despite the differences in absolute pitch height among the various syllables of the low-falling tone, the F_0 curves generally exhibit a gradual decline or flat contour over the syllable, as shown in Figure 5.

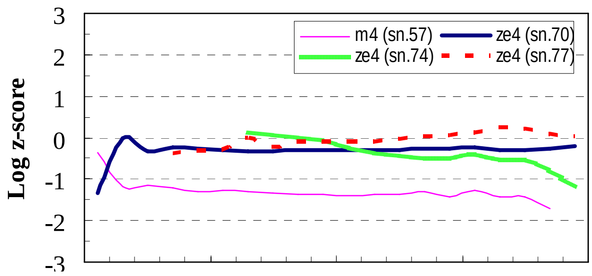

Figure 5. *F_0 curves of the falling tone*

4. Possible problems with the current work

Syllables analyzed in the current study were segmented from utterances in connected speech, which might not provide the most accurate account on the tonal patterns as the contours and values of F_0 could be affected by the F_0 of neighboring syllables due to anticipatory and carryover effects [12].

Besides, unlike adults' speech where intonation comes into most effect on the last few syllables , utterances of young children usually contain fewer syllables (sometimes as few as two to three only in an utterance), thus it proposes the possibility that stronger effect of intonation must have been imposed on the non-final syllables in an utterance.

The above-mentioned effects could have affected the tonal contours of the syllables in the sense that they differ from their counterparts in citation form.

5. Conclusions

Our study demonstrates that the three basic tonal contours (i.e. level, rising and falling) of the respective tones could be generally manifested by the child to a certain extent at this age. However, from the available acoustical data, it seems that within the same type of tonal contour, patterns of production among the different tones are not adequately clear enough for distinction, especially for the three level tones which share the same level tonal contour but being clustered in a narrow F_0 range. Hence, it puts doubt to the claim on early mastery of tones in children acquiring a tonal language.

Furthermore, unlike what has been found in Lee [6]'s study on older children, the production of tone in a thirty-one months seemed not reach steady production yet, in the sense of considerable degree of variability on the patterns of F_0 curves.

6. Acknowledgements

We would like to express special thanks to Thomas Lee for advice and permission for access to the original audio files of CANCORP on behalf of the team. We wish to thank Ying Wai Wong for discussions and comments on this work. We also thank Barbara Lee and other members of the Language Acquisition Laboratory, Language Engineering Laboratory and HNCELA for suggestions and support.

7. References

[1] Bauer, R.S., and Benedict, P.K., 1997. *Modern Cantonese Phonology*. New York: Mouton de Gruyter.

[2] Chao, Y. R., 1947. *Cantonese primer*. London: Oxford University Press.

[3] Chao, Y. R., 1973. The Cantian idiolect: an analysis of the Chinese spoken by a twenty-eight-month-old child. In C. A. Ferguson and D. I. Slobin, eds., *Studies of Child Language*. New York: Holt, Rinehart and Winston.

[4] Hashimoto, A. H., 1972. *Studies in Yue dialect*. London: Cambridge University Press.

[5] Lee, T, H.-T. and Wong, C., 1998. "CANCORP: the Hong Kong Cantonese Child Language Corpus", *Cahiers de Linguistique Asie Orientale vol. 27*, no. 2, pp. 211-228.

[6] Lee, W.S., 1999. *A phonetic study of the speech of Cantonese-speaking children in Hong Kong*. PhD dissertation. The City University of Hong Kong.

[7] Li, C. N. and Thompson, S. A., 1977. The acquisition of tone in Mandarin-speaking children. *Journal of Child Language, 4*, 185-199.

[8] Tse, J. K.-P., 1978. Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language, 5*, 191-204.

[9] Tse, A. C.-Y., 1991. *The acquisition process of Cantonese-phonology – a case study*. MPhil thesis. The University of Hong Kong.

[10] Tuaycharoen, P., 1977. *The phonetic and phonological development of a Thai baby: from early communicative interaction to speech*. Doctoral dissertation, University of London.

[11] Westermeyer, R., and Westermeyer, J., 1977. Tonal language acquisition among Lao children. *Anthropological Linguistics, 19*, 260-264.

[12] Xu, Y., 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics, 25*, 61-83.

[13] Zhu, Z. N., 1999. *Shanghai Tonetics.* München: Lincom Europa.