A Study of Tone Acquisition in a Mandarin-speaking Child under Three

Huiqin Ma, Xiaoxiang Chen, Xin Liu
Hunan University, Hunan Province, China
sophycxx@yahoo.cn

Abstract
How children acquire their L1 has been predicted and explained under different hypotheses. One is the independence hypothesis [1] and the other is the hypothesis of early interaction [2]. The former claimed that L1 development is constrained by biological maturational processes, with no link to perceptual mechanisms, that is, language development is independent of linguistic environment. The latter was that perceptuo-motor attunements already operate in L1 development, so the ambient and environmental effects can be traced from an early age. Our study attempts to explore which perspective is involved in the process. We conduct an investigation on the acquisition of tone in early language development by using longitudinal data of one Mandarin-speaking child living in Changsha. The tonal acquisition process was analyzed from statistical and phonetic aspects to investigate whether Mandarin-learning infants’ tonal development reflect universal or language-specific effects. Our findings reveal that the infant gains a full mastery of tone system of Mandarin till she grew to two years old, or even four months later. Also, as the subject gets more exposure to her mother tongue, her tone production gets approaching to adults’, which indicates the effect of ambient language. All these lend adequate support for the hypothesis of early interaction.

Index Terms: tone acquisition, children under three, southern Mandarin, the hypothesis of early interaction

1. Introduction
When it comes to phonetic and phonological acquisition, suprasegmental features are generally recognized to be acquired earlier than segmental features (i.e. consonants and vowels). Most research on infants’ phonetic and phonological development concentrated on segmental features, while little attention was paid to the prosodic development of infants especially for languages other than English. Even for the small number of investigations conducted on tones, they are mostly on pitch perception at the babbling stage and the acquisition of the tone inventory in early stage of language acquisition. For the few studies on pitch contour production at the babbling stage, they are limited to intonation like [3]. In the study of tone pattern (Japanese) available from [4], the data were not longitudinal and acoustic analysis.

The most salient result on tonal acquisition study is that tones are acquired before segmental inventories and the rising tone is acquired last. Previous Mandarin and Cantonese studies agree that children have more or less mastered the tones at a stage when segments are still quite far from adult forms [5, 6, 7]. In the field of tonal production at the babbling stage, babbling was often observed to have adult-like variable pitch contours, but there are little experimental data in support of or in refutation of the anecdotal view.

Moreover, the issue of independent and interactional development in children’s phonetic acquisition has been addressed in a huge bulk of empirical studies on the acquisition of segmentals, but little is known about suprasegmentals, such as tone in Mandarin. Whether tone in the infant’s vocalizations shares some, if any, properties in the one-word stage and beyond? What kinds of mechanism are involved when the child is approaching the adult tonal system? The present study is an attempt to answer these questions. We intend to investigate the acquisition of tone in early language development by using longitudinal data of one Mandarin-speaking child living in Changsha.

Since the study mainly focuses on the issue of Nature or Nurture especially on verifying the independence hypothesis [1] and the hypothesis of early interaction [2], the research questions for the present study are designed as follows:

(1) At what age does the Mandarin-speaking child acquire the Mandarin tones?
(2) Of the four tones, which tone is acquired first? Is the acquisition of the rising and dipping tones more challenging for the Mandarin-speaking child under three?
(3) Does the infant’s Mandarin tone development support the independence hypothesis or the hypothesis of early interaction?

2. Method
2.1 Subject
The subject, P, was born in Changsha. At home, her parents communicated with her in Southern Mandarin. Both of them are teachers in university. The observation began when the child was one year old and ended when she was two years four months at a monthly interval. The recognizable words/utterances in the 14-hour speech data of the subject were transcribed into IPA and five-point-scale tone letters both by perception and by acoustic methods were labeled. The tonal acquisition process was analyzed from statistical and phonetic aspects to investigate whether the Mandarin-learning infant’s tonal development reflect universal or language-specific effects. The data ranged from 01:00 to 02:04 with 14 hours of audio recordings, including 2294 recognizable utterances, in which 1268 syllables were analyzed.

2.2 Data analysis
The subject was longitudinally observed and her sound samples were recorded in both audio and video tapes for the purpose of a clear understanding of infant speech development. We transferred all the audio recordings to a PC, reserved them into the form ‘wav’, and extracted the utterances of the infant
by dropping her physiological sounds like crying, laughing, creaky and breathy. The utterance is divided according to pause or breathing break. By using the software CoolEdit (Pro.2.0) utterances were extracted from the whole session and CV syllables were extracted by means of the narrow band spectrogram of the speech analysis software Praat from the utterance aided by perception.

Since children’s vocal-cords are still in the process of maturation, and their vocalization may be influenced by different physical and mental states even within the same session, the absolute value of F0 may weaken the comparability of F0 within the same child, hence the necessity of F0 normalization. F0 was normalized for each speaker across four tones. The method of turning normalized F0 values into relative pitch values (T values) according to scale defined by maximum and minimum F0 values is shown in the formula below [9, 10]:

$$T = \frac{\log x - \log (\min - St \min)}{\log (\max + St \max) - \log (\min - St \min)} \times 5$$

(1)

‘x’ is any given point of a pitch contour, ‘min’ represents the minimal data of all measure points, ‘Stmin’ means the standard deviation of the minimal data, and ‘max’ means the maximal data of all, ‘Stmax’ represents the standard deviation of the maximal data. The output (T) is a value from 0 to 5, which is similar to the 5-point pitch scale proposed by [11].

The infant’s tonal patterns were analyzed in terms of accuracy rate with reference to attempted adult patterns. To ensure the reliability of data analysis, we invite three postgraduate classmates majoring in Mandarin Phonetics to accomplish double-check of data. The infant's vocalizations selected for categorization were classified using auditory analysis, supplemented by acoustic analysis using Praat. The intra-reliability was 96.65% and the inter-reliability was 92.5%.

The participant’s performance on tone production was calculated by frequency counts. The accuracy rate was calculated by using the formula below:

$$ACC = \frac{C_i}{O_i} \times 100$$

(2)

ACC, is the percentage of accurate production of each target tones for a certain tonal category; C, is the total number of accurate tone productions, O, is the total number of corresponding attempts.

Generally, there was an agreement that several aspects of fundamental frequency were correlated with tones perceptually: the relative F0 level, the direction of F0 change, and the magnitude of F0 change. [12] proposed three perceptual dimensions labeled “height”, “direction”, and “contour” that were associated with listeners’ perception of tones. He interpreted the “height” as average F0 level, the “direction” as direction of F0 change, and finally the “contour” as the magnitude of F0 change. His “height” dimension and “contour” dimension were similar to our mean F0 and pitch range. The purpose of the analysis was to find to what extent listeners could rely on F0 change and F0 level to identify spontaneously produced tones in infant-directed (ID) speech, and this could shed light on our understanding of the effect of ID speech on child language learning, especially the acquisition of tones.

Referring to the calculation method by [13, 14], accent range can be measured as the logarithmic difference between the maximum and minimum values of the F0 contours and reported in semitones. Measures of semitones were utilized for this portion of the analysis because they are based on a logarithmic scale, and give the best description of our perception of pitch change. The accent range of each non-level pitch pattern is calculated with the following formula:

$$\text{Accent range} = \frac{12}{\log (2)} \times \log \left( \frac{\text{max F0}}{\text{min F0}} \right)$$

3. Results

The 4 figures below respectively depict the developmental trend of mean F0 and accent range of high-level, high-falling, high-rising, fall-rise Mandarin tones.

Figure 1: Statistical descriptions of mean F0 and accent range of high level tone.

The figure above displays the developmental tendency of mean F0 and accent range of high-level tone. Before the child grows to 01:02, the mean F0 and accent range does not develop stably, with some variations among them. Then the child makes great progress and achieves adult-like level at the age of one year one month.

Figure 2 Statistical descriptions of mean F0 and accent range of high falling tone.

Figure 2 illustrates the changing trend of mean F0 and accent range of high falling tone from the infant aging from 01:00 to 02:04. The age period after 01:02:23 sees a general smooth pitch curve and the tone’s mean F0 and accent range grows steadily from the age of 01:03, which indicates that the infant has had a full mastery of this tone category.

Figure 3 Statistical descriptions of mean F0 and accent range of high rising tone.

Figure 3 depicts the development of high rising tone in terms of mean F0 and accent range. Specifically, before the
child was under one year and six months, the two parameters have some changes in the stage, while during the period after 01:06, they develop steadily, which signifies that the infant has mastered this tone at this age.

**Figure 4** Statistical descriptions of mean F0 and accent range of fall rise tone.

Figure 4 shows that the developmental changes of the two values are not as regular as the other three tones. The subject’s mean F0 is irregular before two years and gets stabilized after four months. For accent range, it also keeps changing from 01:00 to 02:00, and becomes more adult-like after four-month progress. There are more fluctuations in the development of this tone than the other three tones across the 14 sessions, which may signify that the infant has more difficulties in producing fall rise tone. She keeps adjusting her production to approach adults’ level during the acquisition process.

These above-mentioned figures show the general results concerning mean F0s and accent ranges of the four tones produced by the Changsha infant. The acquisition process of all the four tones follows the similar pattern, that is, the change of mean F0s and accent ranges keep steady after its acquisition points.

The chart below provides a report of the overall production accuracy and acquisition order of each tone category.

**Figure 5**: Tonal accuracy rate in P’s production across the 14 sessions.

It can be learned from the figure above that high level and high falling tones enjoys a much higher accuracy percentage in the beginning, which signifies its easy acquisition. The subject masters them at 01:01 and high falling 01:02 respectively. The infant, however, takes great efforts to master the other two tones. With the accuracy rates of 21.1% and 15.4% at first, the subject does not make a good performance in the early stage. Even as she grows to 01:02, the infant still feels difficult to produce high rising tone accurately. She finally acquires high rising and fall rise tones after eight-month and four-month progress respectively. Overall, the result is conformity with our acquisition order presented above.

To examine the ambient language effect, we conduct a comparison of F0 features (as represented by T-values) of the four tone types produced by the infant and her caregivers, as shown in the figure below:

**Figure 6**: Comparison in the tone contours of the infant as well as adults.

As is shown in the figure above, when she is one year old, the subject’s tone contours are much deviant from her caregivers’. To be specific, she produces tone 1 as low level, the tone 4 is also immature, and the contours of tone 2 and tone3 are very similar, that is, both of them are articulated as 23 instead of 35 and 214. After two-month progress, Tone 1 and Tone 4 develop greatly to approach to that of adults and achieve adults’ level, which means that she has mastered the two tones at this stage. The subject still has difficulty in distinguishing Tone 2 and Tone 3, suggesting that the child feels much confused about the two tones.

At the age of 01:06, great progress has been made in high rising tone, and the contour and T-value are much similar to her parents’. Also, the dipping tone enjoys great progress in pitch contour. At the age of two years old, the child has an obvious progress in dipping tone, with the other three tones developing more mature. As the subject grows to two years and four months, the contour of tone 2 improves a lot. Specifically, the onset of the contour gets a little steep to approach adults’ pattern.

On the whole, the infant gains improvement with regard to tone contour and T-value of the four tones gradually from the stage of 01:00;10 to 02;04;18. The subject acquires the high level and high falling tones at the age of 01:02, high rising tone at 01:06 and fall rise tone at 02:00 basically and becomes much mature as she grows.

Also, we observe similar durational features in the productions of the infant and her caregivers, as is illustrated below:

**Figure 7**: Comparison in duration of the four tones produced by the infant and her parents.

Figure 7 displays the durational change of the tones studied from the vocalizations of the child and her caregivers. At 01:00, her longest tone is high level tone, followed by fall rise tone and high rising, high falling is the last, which is quite different from the order of her parents in this regard. The situation remains the same until she grows to the age of two years old. At this stage, fall rise tone tops first, followed by
high level and high rising, and high falling is still the last. The following four-month growth helps the subject to develop towards the level of adults greatly, as the figure shows above. Thus, it can be learned that the durational features of the tones are getting more mature as the influence from the ambient language increases.

4. Discussion

The three research questions can be answered as follows:

(1) At what age does the Mandarin-speaking child acquire the Mandarin tones?

Our findings show that there is an order to gain a full mastery of them. By the end of two years and four months, the subject mastered the tone system with some of them not that mature. Statistically, the acquiring point of each tone category differs and large differences in accuracy rate exist around the acquiring period.

(2) Of the four tones, which tone is acquired first? Are the rising and dipping tones more challenging for the Mandarin-speaking child under three to acquire?

By examining the tonal development of the subject, we find that by the age of two years and four months, the child has acquired the high level, high rising as well as high falling tones, but the acquisition of the fall rise tone is not perfect. There is still confusion between high rising and fall rise tones, that is, she often substitutes fall rise tone for high rising tone, which can be observed clearly from the sound pattern charts of the four tones in our acoustic analysis. Our statistical report and maximum and minimum fundamental frequency values also suggest the same findings.

(3) Does the Mandarin-learning infant’s tone production development support the independence hypotheses or the hypothesis of early interaction?

Our findings also reveal a strong support for the hypothesis of early interaction. On the basis of our comparison of tonal features, duration and T-value between the subject and her caregivers, we find that as the subject gets more exposure to her mother tongue, all the values of her tone production get approaching to adults’, which indicates the effect of ambient language. All these lend adequate support for the hypothesis of early interaction.

5. Conclusions

In this paper, we have presented longitudinal data of one Mandarin-speaking child from one year to two years four months to investigate the acquisition of tone in the child’s early speech, to reevaluate previous proposals concerning children’s tonal acquisition and to address two key issues concerning children’s phonological acquisition: the acquisition age and order of tones and the mechanism of tonal acquisition.

6. Acknowledgements

We would like to thank the anonymous reviewers for their guidance and suggestions.

This research entitled “Babbling Vocalizations and Early Speech of Southern Mandarin: A Longitudinal Case Study” has been supported by Plan Projects of Humanities and Social Sciences from the Chinese Ministry of Education [2012]12, by 11ZDB057 from Hunan Province and by Cognitive Science Base of Hunan University.

7. References


