A COMPARISON OF MANDARIN-SPEAKING CHILDREN'S PITCH AT THE PRE-LINGUISTIC AND EARLY SPEECH STAGES

Xiaoling Zhang[a], *Xiaoxiang Chen*[b] *& Huiqin Ma*[b]

[a]National University of Defense Technology, China; [b]Hunan University, China
Congcong96@163.com

ABSTRACT

The paper studies the prosodic development of 9 Mandarin infants (5 males and 4 females) in Changsha through the acoustic and statistic analysis of the infants' data from their pre-linguistic and early speech stages. After the analysis, following results have been obtained: 1) the level and simple prosodic patterns are highly frequent in the infants' prosodic production for the two stages. Furthermore, the infants all prefer the production of high relative pitch. There are similarities as well as differences in the prosodic production between the two age groups. Certain universal characteristics could be observed in comparing the data with those of other language communities; 2) development continuity can be observed between the two age groups in prosodic development; 3) obvious evidence for language-specific effects can be found in both pre-linguistic and early speech stages.

Keywords: infant pitch, continuity, adapting to target language, normalization

1. INTRODUCTION

The data of 9 Mandarin-learning infants (in Changsha) from their pre-linguistic stage (mean age = 0;10;23) and early speech stage (mean age = 1;8;14) were analyzed in my research. We studied the prosodic development by measuring variations in fundamental frequency and relative pitch level rather than investigating tone acquisition, because tone distinguished meaning requires the knowledge of the target language, witch is impossible in the case of pre-linguistic stage. The present research will try to explore the degree of continuity between pre-linguistic and early speech stages and to investigate whether Mandarin-learning infants' pitch development reflected universal or language-specific effect. The aims for present study are as follows.

1. What are the pitch patterns for Mandarin-learning infant in their pre-linguistic and linguistic stages? Are they identical or different? Are there any changes in the distribution of different pitch patterns in different age groups?

2. Are prosodic patterns in Mandarin-learning infants similar to those in other language groups reflecting universal patterns or do these patterns reflect the characteristics of ambient language?

3. Does Mandarin-learning infants' pitch production development support the continuity model?

Table 1: Subjects' demographic data.

Child	Gender	Age Group 1	Age Group 2
TKX	M	0;11;05	**1;08;00**
XJZ	M	0;10;11	**1;08;12**
XWT	M	0;11;08	**1;08;19**
LXM	M	0;11;00	**1;08;16**
ZHZ	M	0;11;20	**1;08;22**
LJ	F	0;11;11	**1;08;14**
HX	F	0;10;17	**1;08;19**
LYL	F	0;10;11	**1;08;12**
TSQ	F	0;10;22	**1;08;13**
Mean	-------	0;10;23	**1;08;14**
SD	-------	**.47**	**.21**

2. METHODS

2.1. Participants

Two sessions of each subject were selected, one is from the pre-linguistic stage, and the other is from the early speech stage. The infants (see Table 1) met the following criteria: 1) no unusual parental, sensory, or developmental concerns; 2) from Mandarin-speaking homes; 3) normal hearing limits defined as 20 dB HL or better at 500, 1000, 2000 Hz; 4) the first session of each subject was recorded around 10 to 11 months, with a mean age of 0;10;23 and the second session was recorded at around 1;8;0 old, with a mean age of 1;8;14. The age of the subjects in each age group was limited within 1 month difference with the mean one. Standard Derivation of age calculated by SPSS was .47 and .21 respectively for Age Group 1 and Age Group 2, which means the infants' age in each age group do not have big difference. Therefore, they could be listed in the same age group. They will be called Age Group 1 (*pre-linguistic group*) and Age Group 2 (*early speech stages*) in the following research.

2.2. Recording procedures

During the recording, the infants are exposed in a natural setting with the surrounding people. Care was taken to create a natural atmosphere of play or verbal communication. The observers and the infants' family members like mother, father, or grandmother attended the recording process. The infants were familiar with the observers. We elicited the infants' production by playing materials such as blocks, balls, doll baby, animal puppet, and picture books. During the recording process, we turned off all the electrical appliances so as to exclude the electrical interference in the recording. Their sound samples were recorded both by sound recording pen (SONY ICD-SX35) and video camera (SONY HDR-HC1) for the purpose of a clear understanding of the development of infant speech. Each recording was from 45 to 60 minutes. All the data are from Hunan Infant Children Speech Database.

2.3. Data analysis

That data collected in the 18 sessions will be extracted, segmented, converted and normalized according to the following ways.

2.3.1. Extraction and segmentation

We reserved the audio recordings into the form 'wav', extracted the utterances of the children and dropped the children's physiological sounds like crying, laughing, creaky and breathy. Each utterance included the babbling excluding the noises. The utterance was divided according to pause or breathing break. CV syllables were extracted by means of the narrow band spectrogram of the speech analysis software PRAAT from the utterance aided by perception. The requirements of extracting of CV s are referred to Chen 2005 [3].

2.3.2. Conversion and normalization

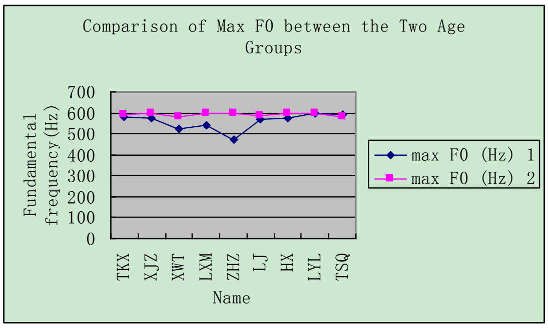
Individual difference may greatly affect the description of pitch and tone and the fundamental frequency of a speaker is basically a physiological determined characteristic and is speaker-dependent. Therefore, when we do the comparison of pitch, we need to conduct the process of normalization for fundamental frequency. The pitch contour of each CV syllable was processed by PRAAT, with eleven F0 values chosen for each continuous contour at equally spaced intervals. All F0 values were normalized by Lz-score within each session and the Lz-score value of each CV was converted into a relative pitch level description on a five-point scale. In this research, I referred to Zhu's [12] research to normalize F0 values of the infants' utterances. Zhu defined the pitch range with logarithmic mean value and standard deviation. The formula of Lz-score calculation is as follows.

$$(1)\quad z_i' = \frac{y_i - m_y}{s_y} = \frac{\lg x_i - \frac{1}{n}\sum_{i=1}^{n}\lg x_i}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\lg x_i - \frac{1}{n}\sum_{i=1}^{n}\lg x_i)^2}}$$

In the formula, x_i is the absolute fundamental frequency of the sample point. And y_i is the logarithmic value of fundamental frequency of x_i. m_y and s_y are respectively the arithmetic mean value and standard deviation of y_i (i=1,2…,n). m_y is logarithmic mean value of the original fundamental frequency. z_i gives the Lz-score of y_i, which measures the distance of y_i from the mean in terms of the number of standard deviations. In the view of statistics, the logarithmic value of fundamental frequency approximates to the distribution of the perception of pitch and the fundamental frequency than the absolute value of fundamental frequency.

The five-point scale was used and the 'tone letter' notation proposed by Y. R. Chao [2] was adopted in tonal transcription of my research.

Figure 1:.Comparison of maximum F0 between the two age groups.



3. RESULTS

3.1. Maximum, minimum and mean F0

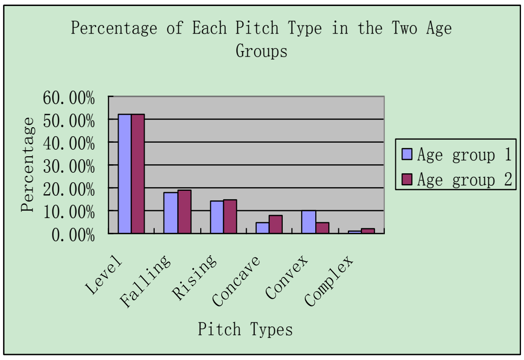
The maximum F0 ranged from 500Hz to 600Hz and the minimum F0 ranged from 76Hz to 120Hz with no significant group differences. Wilcoxon tests were used to analyze the differences of max-F0 and min-F0 between two stages. Z = -2.192, p =.028, so there was significant difference between the two age groups in terms of maximum F0 and there was an obvious increase of maximum F0 values for Age Group 2 (Figure 1). To test minimum F0 between the groups, Wilcoxon tests were adopted, p = .214 (> 0.05), therefore no significant difference was found. And there is no significant difference of mean F0 values between the two age groups, either.

3.2. Analyses of pitch types

After converting the pitch contours into five-point scale, the relative pitch values could generally be divided into the following types: level, falling, rising, concave, convex and multiplex pitch. This categorization was based on the research of Cruttenden [5] and Balog's [1] research.

In the total 18 sessions, simple pitch types like level, falling and rising were the main constituents of the infants' pitch production in both age groups. Level pitch predominated in the pitch production in both of the two age groups. The least frequently occurring pitch type was complex contour, which totaled only 1.6% in our data, a very small part, so we ignored them in the research.

Figure 2: Comparison of the Level, Falling, Rising, Concave, Convex and Complex in the Two Age Groups



Level pitch was the most predominant pitch contours, contributing to over 50% (Figure 2) of the total number of pitch types in both age groups. Both falling and rising are non-level pitch, in which falling pitch was the next frequent one, followed by rising pitch. The percentage of falling and rising pitch contours in Age Group 2 all increased comparing with that of Age Group 1, however no significant difference was found by Wilcoxon test (p=0.59; 0.31). The number of convex pitch is higher than that of concave pitch in Age Group 1. The percentage of concave and convex in two age groups is 4.7%:10%. Significant differences of concave and convex distribution in the two age groups, could be detected (p = .021; .05).

3.2.1. Level pitch

According to the five-point scale, level pitch could be divided into 3 types: high level (55, 44), mid level (33) and low level (11, 22). It was found that high level pitch (55, 44) was predominant among all level contours, which consisted of 47.0% and 47.1% of the total number of pitch contours respectively in Age Group 1 and Age Group 2. And in the high level pitch patterns, the number of 44 pitch pattern is the most frequent one, while the low level pitch (11, 22) is the least frequent one in both of the two age groups. The distribution of mid level pitch for the two age groups was respectively 4.9% and 4.7%, constituting a very small part. From this we can see that infants' pitch preferred high level pitch than others. It showed no significant difference on this point in late pre-word stage and early speech stage.

3.2.2. Falling and rising contours

Few sharp rises could be found among all the falling and rising pitch contours. The frequency of 54 and 43 was very high, while that of 52, 51 was very low. It seemed that most of the subjects preferred high falling pitches like 54, 53 and 43, and didn't prefer mid fall and low fall pitches like 31 and 21. High rising 45 and mid rising (35, 34) contours were the most frequent in the rising pitch contours, while the low rise patterns like 25, 12, 13, 15 were the least frequent ones. High rise and mid rise patterns constituted the majority of the rising pitch contours.

4. DISCUSSION

The Mandarin-learning infants in Changsha shared the preference of level and simple prosodic patterns and high relative pitch in the two stages. The frequency of level pitch pattern and high relative pitch predominated in both of the pre-linguistic and early speech stages. These similarities in prosodic development in the two stages reflected the continuum tendency.

Compared with other researches on the prosodic development, universal characteristics could be observed. Level contour was also found most commonly produced in a cross-sectional study of 20 infants aged 0;3;0, 0;6;0 and 0;9;0 by Kent and Murray's study [7], and second commonly found during the first two years of life by the research of Robb [10]. Concave, convex and other complex contours were rarely observed. In the simple pitch patterns, the rises and falls without sharp changing degree were favored by the infants in the two stages. These findings provide evidence that pre-word stage is governed by general restrictions of human phonological capacity.

On the other hand, the infants' rich production of level, simple pitch patterns and high relative pitch and poor production of concave, convex and other complex contours consisted with the characteristics of Chinese according to Cheng's [4] quantitative research which showed that high tones have predominated in most dialects and the largest number of tones is either level tones or falling tones, while bidirectional tone like concave and convex tones have the least number of occurrence. This is an obvious evidence for the language-specific effects on the infants' prosodic development. Moreover, the infants' production of four main Mandarin tones: high-level, high-rising, low-falling-rising, and high-falling in the two age groups all surpass 50%. All these facts displayed the effects of ambient language.

There are some differences on the prosodic development between the two age groups. First, the frequency occurrence of the predominant prosodic pattern level pitch in pre-word stage decreased slightly in the early speech stage and the portion of simple contours increased correspondingly. The change balanced the distribution of fundamental frequency contours in early speech stage. The decrease of level pitch pattern and the increase of falling and rising pitch contours from pre-word to the early speech stages might be a developmental trend reflecting the influence of the language environment. Secondly, significant differences in the distribution of concave and convex contours in the two age groups may cause by the interaction between perception and production [9, 11].

5. CONCLUSION

In conclusion, we found no evidence of a sharp discontinuity postulated by Jakobson [6] between the pre-linguistic and early linguistic stages. On the contrary, our quantitative measurements and analysis revealed significant similarities between the two age groups. In addition, the differences between the two age groups were part of a continuum of developmental changes that gradually brought infants' vocalizations closer to the adult model, similar to the first and third continuous stages in the neurobiological approach [8].

6. REFERENCES

[1] Balog, H.L., Snow, D. 2007. The adaptation and application of relational and independent analyses for intonation production in young children. *Journal of Phonetics* 35, 118-133.

[2] Chao, Y.R. 1930. A system of tone letters. *La Maitre Phonetique* 45, 24-27.

[3] Chen Y. 2005. Characteristics of pitch at the babbling dtage. *Journal of Lingling University* 1, 183-185.

[4] Cheng, C.C. 1973. A quantitative study of Chinese tones. *Journal of Chinese Linguistics* 1, 93-110.

[5] Cruttenden, A. 1997. *Intonation* (2^{nd} ed.). Cambridge: Cambridge University Press.

[6] Jakobson, R. 1941/1968. *Child Language, Aphasia, and Phonological Universals*. Translated by Keiler, A.R. The Hague: Mouton.

[7] Kent, R.D., Bauer, H.R. 1985. Vocalization of one-year-olds. *Journal of Child Language* 12, 491-526.

[8] Locke, J.L. 1983. *Phonological Acquisition and Change*. New York: Academic Press.

[9] Locke, J.L., Pearson, D.M. 1992. Vocal learning and the emergence of phonological capacity. In Ferguson C.A, Menn, L., Stole-Gammon., C. (eds.), *Phonological Development: Models, Research, Implications.* Timonium: York Press.

[10] Robb, M.P., Saxman, J.H., Grant. A.A. 1989. Vocal fundamental frequency characteristics during the first two years of life. *Journal of the Acoustical Society of America* 85, 1708-1717.

[11] Vihman, M.M. 1993. Variable paths to early word production. *Journal of Phonetics* 21, 61-82.

[12] Zhu X.N. 2004. F0 normalization: How to deal with between-speaker tonal variations? *Linguistic Sciences* 2, 3-19.