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

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

Lexical variation and rime-tone correlation in early tonal acquisition: a longitudinal study of Mandarin Chinese*

Jie Yang[1] & *Thomas Hun-tak Lee*[2,1]

Institute of Cognitive Science, Hunan University, Changsha[1]
Language Acquisition Laboratory, Chinese University of Hong Kong, Hong Kong[2]

thyjessie@yahoo.com.cn, Huntaklee@cuhk.edu.hk

1. Introduction

Studies of the acquisition of tones in Chinese dialects have documented the early mastery of tones in terms of accuracy in children's tonal production, suggesting that children produce the distinctive tones of their target language competently by age two or earlier (Li and Thompson 1977, J. Tse 1977, Clumeck 1980, Hsu 1996, Zhu and Dodd 2000). These studies have also revealed interesting patterns of order of tone acquisition (such as the level tone and the high falling tone occurring before the fall-rise tone in Mandarin), and of tonal substitution in child language (such as the substitution of a high rising tone for the fall-rise tone in Mandarin). However, the claim about early tonal acquisition deserves critical scrutiny even at the level of phonetic production, given that detailed longitudinal studies have revealed that some tones such as the low rising in Cantonese are acquired late (A. Tse 1993), and a clear lexical effect has been observed for the acquisition of Taiwanese tone (Tsay 2004). In line with other studies suggesting lexical diffusion in phonological development (Ferguson and Farwell 1975, Hsieh 1977), the fact that children can produce a tone accurately on selected morphemes may not mean s/he can produce the same tone accurately for most or all lexical items being acquired.

Recent pitch studies of the babbling of infants acquiring Cantonese and Mandarin have revealed preferences for level pitch over falling and rising pitch contours, which were in turn favored over concave and convex contours (Chen 2005, Lee, Lee and Chen 2005). How these preferential pitch patterns in babbling evolve into the adult tonal system would need to be understood on the basis of acoustical analysis of infant pitch production, as well as analysis of adult input.

The present study addresses the following issues: (a) What is the acquisition order of Mandarin tones, with respect to consistency and accuracy of tonal production? (b) To what extent does accuracy of tonal production depend on the lexical dimension? (c) Given that tone needs to be realized on the syllable, would there be any correlation between tonal complexity and the complexity of the rime of the syllable, in the spirit of Woo (1969) and Yip (1980,1989) (d) Do the tonal substitution patterns in child language reflect early preferences for pitch contour types?

2. Data and method

2.1. Data collection

Hour-long audio- and audio-visual recordings of a Changsha child (AJR) acquiring Southern Mandarin as her native language, from 1 year 14 days to 2 year 3 months and 13 days, were used for analysis. The child was observed interacting with the parent and investigators in naturalistic settings, at weekly intervals up to two years of age, and semi-monthly after the child reached two. The audio recordings, made on a SONY TCS-100 cassette recorder, were digitalized into wave files for segmentation and analysis. Altogether 18 sessions were included.

2.2. Data segmentation

The wave forms of each session were segmented and recognizable words were extracted into separate files. Based on Vihman (1996), a recognizable word was either (i) a form recognized as a word by caretaker(s) involved in the interaction, i.e. adult repetition of child's vocalization with the corresponding adult form; or (ii) a form suggested by the context, i.e. the context strongly suggested the corresponding adult form of the child's vocalization and no other. Among the recognized vocalizations, unclear words, exclamatory words, and onomatopoeic words were not included. Both spontaneous and imitated vocalizations were admitted.

The extracted recognizable words were segmented into syllables by PRAAT for the convenience of phonetic transcription. The beginning and end points of each pitch curve generated by PRAAT were noted, but only the portion clearly reflecting the pitch contour was selected for sampling of F0 values. Eleven equally spaced points of each pitch contour were selected for determination of relative pitch contour shape.

2.3. Transcription

The phonetic segments of each word were transcribed impressionistically in IPA symbols, and checked against spectrograms. The tones appearing in each word were determined by a normalization procedure converting the 11 F0 values into relative pitch on the five-point scale of Chao (1930). These were checked against impressionistic perception by one of the investigators.

The method of Zhu (2004) for normalizing F0 values was adopted for conversion from F0 to relative pitch. In employing this method, we used the mean of log z scores of F0 values from each child recording session as defining the middle of the relative pitch range for that session. On a five-point scale, a standard deviation (s.d.) of +1.5 or more will count as point 5; a s.d. of +0.5 to +1.5 will be considered as point 4; a s.d. range of -0.5 to +0.5 covers point 3; a s.d. of -0.5 to -1.5 defines point 2, and a s.d. of -1.5 and lower covers point 1.

3. Results

3.1. The production accuracy of Mandarin tones

Mandarin has four distinctive tones: Tone 1 (a high level , 55), Tone 2 (a high rising, 35), Tone 3 (a fall-rise, 214), and Tone 4 (a high falling, 51) (Chao 1968). A pitch contour was considered as produced accurately by AJR when it matched the tonal contour of the adult target. All low falling contours AJR produced for the realization of Tone 3 target words were also considered as accurate. The calculation of accuracy rate followed the formula below:

$$ACC_i = \frac{C_i}{O_i} * 100 \quad (1)$$

where ACC_i is the percentage of tones accurately produced for a certain tonal category; C_i is the number of tokens of the tone produced correctly, O_i is the number of opportunities for the production of the tone.

Only words produced in utterance-final position and in isolation were used in this analysis in order to avoid possible effects of tonal interaction in other positions. Since the accuracy rate on isolated syllables and those on utterance-final syllables did not differ significantly with respect to Tone 1, Tone 3 and Tone 4, the two sets of data were pooled together. The results on production accuracy were plotted over 18 sessions, as in Figure 1.

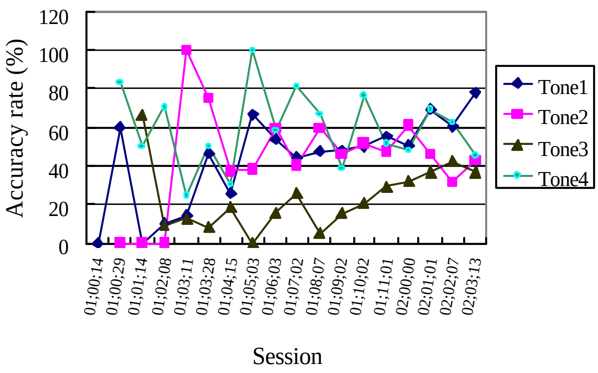

Figure 1: *Tonal accuracy in AJR's production.*

Unlike findings in previous studies indicating that Tone 1 and Tone 4 were acquired before Tone 2 (Li and Thompson 1977, Clumeck 1980), Tone 2 was produced at an accuracy rate comparable to Tone 1 in the present study. Tone 3 had a low accuracy rate of below 40% throughout the period studied, confirming its later acquisition. However, if an overall accuracy rate of 70 % was adopted to evaluate the mastery of a tone category, three of the four Mandarin tones had not been acquired at the end of the observation period. This conflicts with the claim of the early acquisition of tonal systems, but is in accord with the recent study of Wong, Schwartz & Jenkins (2005), who showed that tone production by three-year-old Mandarin-speaking children in monosyllabic words was not yet adult-like.

3.2. Lexical variation in tonal development

In this section, morphemes which appeared three or more times in utterance-final position and in isolation were included for analysis. Given the monosyllabicity of Mandarin Chinese, each morpheme in this study corresponded to one syllable, and in most cases also corresponded to a word or the first or second syllable of a disyllabic word. The rate of accuracy of tonal production was calculated for each morpheme in each session. Those morphemes produced with an accuracy rate of 70% or higher were considered as accurately pronounced. The age at which AJR produced a target tone accurately on three or more morphemes was taken as the point of acquisition of that tonal category. By this lexically-based criterion, as illustrated in Table 1, Tone 4 was acquired first, followed by Tone 1 and Tone 2, with Tone 3 being the last to be acquired.

Table 1: *Accuracy of tonal production on morphemes.*

Tone category	a	b	c
Tone 1 (High level)	01;06;03	9	5
Tone 2 (High rising)	01;06;03	8	7
Tone 3 (Fall-rise)	02;01;01	10	5
Tone 4 (High falling)	01;02;08	4	4

a= Age at which the tone was produced with at least 70% accuracy on three or more morphemes (Session A); b= Number of morphemes bearing the tone attempted by AJR at Session A; c= Number of morphemes bearing the tone pronounced with at least 70% accuracy at Session A.

From Table 1, we observe that when AJR had reached the age when a tone category was acquired, some morphemes in that tone category still showed an accuracy rate of lower than 70%. For Tone 1, of the nine morphemes attempted, only five were accurately produced at the age when production accuracy reached criterion. For Tone 3, in the session when AJR could accurately produce the tone on three or more morphemes, only half of the morphemes attempted were accurately produced. The table indicates clearly that for three of the four Mandarin tones, AJR was not able to produce a tone accurately for all morphemes bearing the tone in her lexicon even at the point of acquisition.

By analyzing the accuracy of each morpheme across the observation period, we find that for the same tone category, when some lexical items had reached a high accuracy rate at the end of the observation period, others were still not well articulated with respect to tone. In the four tones, all the number words showed a higher accuracy rate than other words. There was no correlation between the number of lexical tokens and tonal accuracy. It was not the case that the most frequent lexical items reflected a high accuracy rate and the least frequent lexical items a low accuracy rate. In the four tone categories, the lexical items with the top three accuracy rates were not the most frequently occurring ones.

Apart from practicing the vocal ability to control the pitch contours of Mandarin Chinese, the child acquiring Mandarin tones needs to understand that tone is contrastive and establish tonemic contrasts. In the present analysis, following the method of Ferguson & Farwell (1975) and Stoel-Gammon (1984) for analyzing the lexical factor in segmental acquisition, tone trees were built for each tone category, including the three morphemes with the highest and lowest accuracy rates in that tone category, as in Figure 2. The lexical items listed at the top of the figure were arranged in decreasing order of accuracy from left to right. The number in the parenthesis beside a morpheme represents the total number of its tokens in the sessions concerned. All the realizations of the target tone for each morpheme are listed in the square boxes with the number indicating the number of tokens. L, R, F, lf, cc, cv refer to level, rising, high-falling, low-falling, concave and convex contours respectively. The leftmost column indicates the ages at which the tokens of the morphemes appeared.

The six morphemes included for each tone illustrate two types of variation. First, a morpheme bearing a certain target tone could be accurately produced when other morphemes bearing the same tone were not, reinforcing the finding in Table 1.

Take the morphemes *ba1* "eight" and *wa1* "frog" for Tone 1 in Figure 2A. These were produced accurately in 1 year 6 months and 3 days, but the morphemes *zhen1* "character in name" and *ban1* "job" were not accurately rendered even by 1 year 7 months 2 days. Consider also Figure 2C for Tone 3, which can be rendered as a fall-rise or a low falling, the child had produced two morphemes on this tone accurately by 1 year 7 months 2 days (*san3* "umbrella" and *ba3* "shit"). However, the same tone falling on the morphemes *shui3* "water" was not accurately produced even as late as 1 year 11 months.

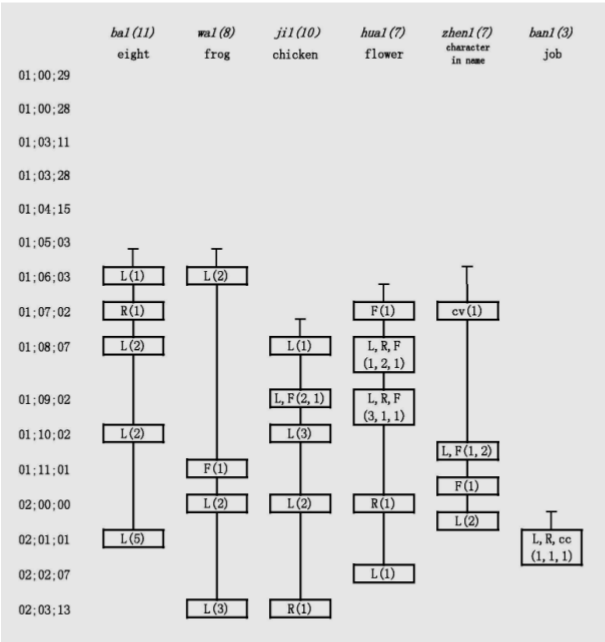

Figure 2A: *Tone tree showing the tokens of the three most accurately pronounced morphemes and the three least accurately pronounced morphemes for Tone 1 (high level).*

Secondly, the tone trees document the variation found in tonal production when we trace the pronunciation of a morpheme over time. In the case of both *wa1* "frog" and *ba1* "eight" for Tone 1, the morpheme was produced as a non-level tone even after its accurate production as a level tone at an earlier session. Consider the Tone 2 morpheme *pa2* "climb" in Figure 2B. It was inaccurately pronounced as a level tone or as a falling tone at the beginning, and was first correctly rendered as a rising tone at 2 years 1 month and 1 day. But subsequently, two months later, at 2 years 3 months, the child reverted to a level tone when using this morpheme. Both types of variation can be found for Tone 4, a high falling tone, as can be seen from Figure 2D.

The longitudinal data from AJR show clearly that the acquisition of Mandarin tone is not an across-the-board phonological process, but one that is intimately linked to the development of children's lexical repertoire. This confirms the earlier findings of Ferguson and Farwell (1975), Hsieh (1977), Stoel-Gammon (1984), and the recent study of Tsay (2004).

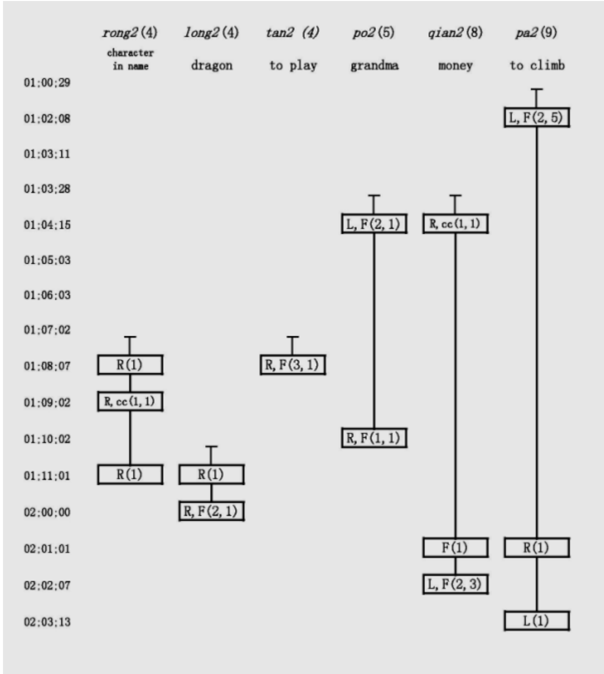

Figure 2B: *Tone tree showing the tokens of the three most accurately pronounced morphemes and the three least accurately pronounced morphemes for Tone 2 (high rising).*

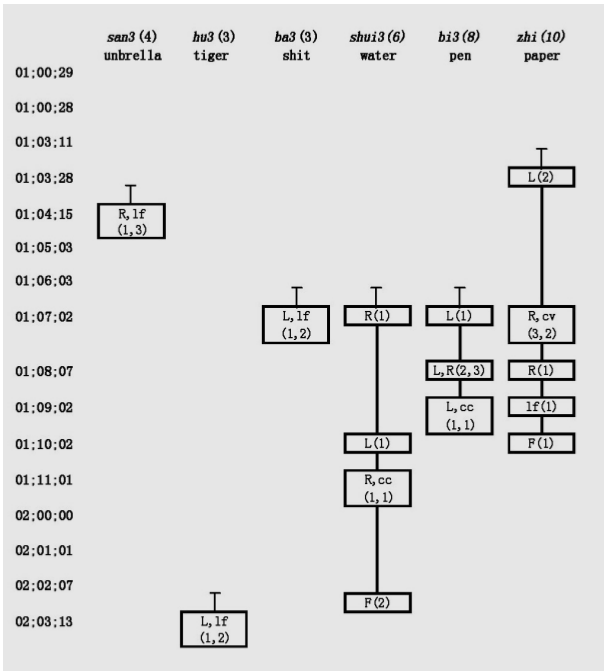

Figure 2C: *Tone tree showing the tokens of the three most accurately pronounced morphemes and the three least accurately pronounced morphemes for Tone 3 (fall-rise).*

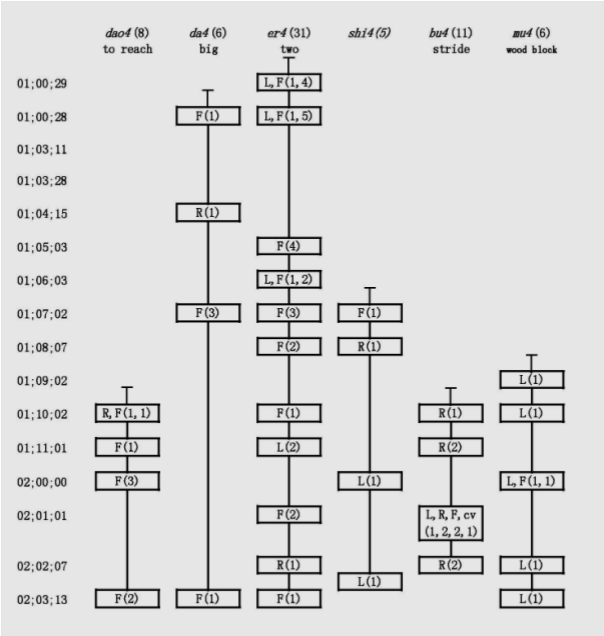

Figure 2D: *Tone tree showing the tokens of the three most accurately pronounced morphemes and the three least accurately pronounced morphemes for Tone 4 (high falling).*

3.3. Correlation between syllable complexity and tone contour complexity

In this section, the correlation between the complexity of the syllable rime and that of the tone contour is examined. The tonal contours of syllables appearing in all positions in an utterance were included in this analysis. In this paper, we consider rising tones and falling tones as simple contour tones, and fall-rises (concave tone) and rise-falls (convex tone) as complex contour tones. The rime of a syllable is considered as complex if it consists of a diphthong (VV) or a VN sequence. We found that complex contour tones had a stronger tendency to fall on complex rimes than the level tone and simple contour tones. This tendency was observed in AJR's phonetic production and in her tonal errors.

3.3.1. Syllable and tone in AJR's tonal production

As shown in Table 2, AJR was inclined to use complex rime for the realization of complex contour tones, i.e. concave and convex tones, and simple rime for the realization of non-complex tones, i.e. level, rising and falling tones.

Table 2: *Tones produced by AJR according to complexity of syllable rime (token)*

Tones	Simple rime		Complex rime	
	token	%	token	%
Level	1627	77.96	460	22.04
Rising	931	65.24	496	34.76
Falling	1203	72.3	461	27.7
Complex contour (concave & convex)	68	26.36	190	73.64

3.3.2. Syllable and tone in AJR's tonal errors

The correlation between syllable complexity and contour complexity was also observed in AJR's tonal errors. As shown in Table 3A, when AJR substituted a complex contour tone (concave or convex) for a non-complex target tone (Tone 1, Tone 2 and Tone 4), and committed a rime error, 65% to 100% of these rime errors involved a change from a simple rime into a complex one. When she replaced a non-complex target tone with another non-complex tone, and made a rime error at the same time, 55% to 89% of the rime errors consisted of a change from a complex rime into a simple one, as shown in Table 3B. Similarly, when she replaced a complex target tone (Tone 3) with a non-complex tone (i.e. a level tone, a rising tone, or a high falling tone), and produced a rime error concurrently, 95% of the rime errors were changed from complex rimes into simple rimes, as shown in Table 3C.

Table 3A: *Tone production errors with non-complex tone replaced by complex tone.*

Target tone	d	e	f	g	h	i	%
Tone 1 (High level)	672	30	10	20	7	7	100
Tone 2 (High rising)	478	105	28	77	11	10	90.91
Tone 4 (High falling)	1005	47	12	35	17	11	64.71

d=Number of tone errors; e=Number of target tones realized as complex tone; f=Number of simple rimes bearing the complex tone; g=Number of complex rimes bearing the complex tone; h=Number of rime errors; i=Number of simple rimes changed to complex rimes; %=Percentage of simple rimes changed into complex rimes.

Table 3B: *Tone production errors with non-complex tone replaced by noncomplex tone other than the target one.*

Target tone	d	e	f	g	h	i	%
Tone 1 (High level)	672	642	501	141	77	42	54.55
Tone 2 (High rising)	478	373	242	131	101	90	89.11
Tone 4 (High falling)	1005	958	686	272	217	162	74.65

d=Number of tone errors; e=Number of target tones realized as other non-target non-complex tone; f=Number of simple rimes bearing the non-complex tone; g=Number of complex rimes bearing the non-complex tone; h=Number of rime errors; i= Number of complex rime changed into simple rime; %=Percentage of complex rime changed into simple rimes.

Table 3C: *Tone production errors with complex tone replaced by non-complex tone.*

Target tone	d	e	f	g	h	i	%
Tone 3 (Fall-rise)	1292	1292	292	800	550	523	95.09

d=Number of tone errors; e=Number of target tones realized as non-complex tone; f=Number of complex rimes bearing the non-complex tone; g=Number of simple rimes bearing the non-complex tone; h=Number of rime errors; i= Number of complex rimes changed to simple rimes; %= Percentage of complex rimes changed into simple rimes.

3.3.3. *Syllable and tone in AJR's syllable errors*

When AJR produced the target tone correctly but erred on complexity of syllable rimes, she would also exhibit the rime-tone correlation as well. As shown in Table 4A, when AJR correctly produced non-complex target tones (Tone 1, Tone 2, and Tone 4), a great majority of the tones were realized on simple rimes. In those instances where a non-complex target tone was correctly produced but a rime error was observed, 64% to 89% of these rime errors involved changes from a complex rime into a simple one. On the other hand, when AJR produced a complex target tone (Tone 3) correctly, the tone occurred mostly on syllables with complex rimes. 80% of the rime errors in the correctly produced Tone 3 tokens consisted of changes from a simple rime into a complex one, as shown in Table 4B.

Table 4A: *Syllable production errors with correctly uttered non-complex tone*

Target tone	j	k	l	m	n	%
Tone 1 (High level)	805	622	183	118	76	64.41
Tone 2 (High rising)	518	310	208	64	57	89.06
Tone 4 (High falling)	427	303	124	112	91	81.25

j=Number of correctly uttered target tone; k=Number of simple rimes bearing the non-complex tone; l=Number of complex rimes bearing the non-complex tone; m=Number of rime errors; n= Number of complex rimes changed into simple rimes; %= Percentage of complex rimes changed into simple rimes.

Table 4B: *Syllable production errors with correctly uttered complex tone*

Target tone	j	k	l	m	n	%
Tone 3 (Fall-rise)	50	10	40	5	4	80

j=Number of correctly uttered Tone 3; k=Number of simple rimes bearing the complex tone; l=Number of complex rimes bearing the complex tone; m=Number of rime error; n= Number of simple rimes changed into complex rimes; %= Percentage of simple rimes changed into complex rimes.

3.3.4. *Syllable and tone in adult input to AJR*

By examining the adult input to AJR, we observe that the above-mentioned association of tone complexity with rime complexity reflects in part an internally driven process of the child, and cannot be entirely attributed to co-occurrence frequencies of syllable rime and tone complexity in the adult input.

The adult utterances in 14 of the recording sessions, separated by monthly intervals, were analyzed, yielding a total of 103,617 Chinese morpheme tokens, classified according to tone category and syllable type. In Table 5, one can see that the complex contour tone (Tone 3) had a slight preference for complex rimes over simple rimes (53% v.s. 47%). Some non-complex tones (Tone 1 and Tone 4) favored simple rimes over complex rimes, with varying degrees of preference. However, it is not the case that all non-complex tones favored simple rimes. For instance, the high rising tone (Tone 2) preferred complex rather than simple rimes, differing from the pattern observed for Tone 1 and Tone 4. Therefore, AJR's tendency to use complex rimes for the realization of complex contour tone, and simple rimes for the realization of non-complex ones, cannot be entirely traced to the adult input. The correlation between the rime complexity of the syllable and pitch contour complexity may have a universal phonetic basis, and provides psycholinguistic support for the decomposition of complex contour tones into simpler ones (Woo 1969; Yip 1980, 1989; Bao 1999).

Table 5: *Tones in the adult input to AJR according to the complexity of syllable rime (token).*

Tones	Simple rime		Complex rime	
	token	%	token	%
Tone 1 (High level)	20469	73.46	7397	26.54
Tone 2 (High rising)	5155	31.2	11370	68.8
Tone 3 (Rise-fall)	11998	46.67	13711	53.33
Tone 4 (High falling)	18965	56.58	14552	43.42

3.4. Tone substitution patterns

Table 6 summarizes the substitution patterns for the four tone categories in AJR's production of utterance-final and isolated syllables. Generally, target Tone 1 (high level) was replaced by a falling or rising tone; target Tone 2 (high rising) appeared mostly as a level tone or a falling tone; target Tone 3 (low fall or fall-rise) was typically produced as a high falling tone or as a level tone; target Tone 4 (high falling) was often replaced by a level or a rising tone. It should be observed that while a convex tone is not a distinctive tone of the Mandarin tonal system, it was used for substituting all tone categories.

Table 6: *Tone substitutions of AJR according to tone category.*

Target tone	Level	Rising	Falling	Concave	Convex
Tone 1	-	49	66	3	12
Tone 2	76	-	65	48	30
Tone 3	79	36	120	-	5
Tone 4	77	58	-	6	21
Total	232	143	251	57	68

The data showed that complex contour tones (concave and convex) were seldom used to substitute for the high level tone (Tone 1), the high falling tone (Tone 4), or the fall-rise tone (Tone 3). However, these complex tones substituted for the high rising tone (Tone 2) to a certain extent. The pattern of tonal substitution reflected a preference for falling and level tones, followed by the rising tone. The least favored category was complex contour tones. This is reminiscent of the preferential patterns found in infant pitch production (Chen 2005, Lee, Lee and Chen 2005).

As for the substitution pattern for target Tone 1, our findings echo those of Wong, Schwartz & Jenkins (2005), who found that half of the Tone 1 tokens were replaced by a falling tone and the other half by a rising tone. The substitution pattern for target Tone 2 resembles that found in Hsu (1996). However, while Tone 1 and Tone 4 were never replaced by a rising tone in Hsu (1996), these tone targets were often replaced by a rising tone in our study.

What is intriguing is the substitution pattern for Tone 3. All previous studies (Chao 1951, Li and Thompson 1977, Clumeck 1980, Zhu & Dodd 2000) claim that Tone 3 was easily confused with Tone 2 in child tonal production. In Hsu (1996), Tone 3 was replaced by a level or falling tone at the early stage of tonal acquisition and by a rising tone at a later stage. The difficulty in mastering the distinction between Tone 2 and Tone 3 was attributed to the phonetic resemblance of the two tones, each having a rising component, and to their alternating with each other in the Tone 3 sandhi environment in Mandarin Chinese. However, in AJR's data, Tone 3 was most frequently replaced by a high falling contour resembling that of Tone 4 rather than by Tone 2. This may be related to the preponderance of Tone 3 tokens realized as a low falling tone in the adult input.

4. Conclusions

In this study, if an overall production accuracy rate of 70% or more is taken as the point of acquisition (see Figure 1), three of the tones had not been acquired even as late as 2 year 3 months. On the other hand, if one considers accurate production on three morphemes as the criterion of acquisition of a tone category, Tone 4 (high falling) was acquired by 1 year 2 months, Tone 1 (high level) and Tone 2 (high rising) reached criterion at 1 year 6 months, while Tone 3 was not firmly grasped until 2 year 1 month. Our study does not point to early tonal acquisition once the lexical factor is taken into account.

The longitudinal data on tonal production indicate convincingly that tonal acquisition is not an across-the-board phenomenon but diffuses through the lexicon, highlighting the earlier insights of Ferguson and Farwell (1975) and Hsieh (1977) and agrees in spirit with the recent findings of Tsay (2004). Tonal accuracy on one morpheme does not imply tonal accuracy on another morpheme bearing the same tone. Tonal accuracy on one instance of a morpheme may not always imply consistent performance on the same morpheme at the beginning stages of tonal development.

Our findings reveal a strong tendency for the child to use complex contour tones on complex rimes, and level, rising and falling tones on simple rimes, as evidenced from the child's tonal production and her tonal errors. This tendency cannot be solely attributed to input influence, and is indicative of an intrinsic association of supra-segmental features with segmental material, lending support to the decompositionality of complex contour tones (Woo 1969; Yip 1980, 1989; Bao 1999).

The tonal substitution patterns reflect a preference for the falling tone and the level tone, followed by the rising tone, over concave and convex tones. This echoes the pitch pattern preferences found in infant babbling (Chen 2005, Lee, Lee and Chen 2005). A prominent error for Tone 3 was substitution by a falling tone rather than a rising tone.

5. References

[1] Bao, Z. M., 1999. The structure of tone. New York: Oxford University Press.

[2] Chao, Y. R., 1930. A system of tone letters. Reprinted in *Fangyan* 1980 (2), 81-83.

[3] Chao, Y. R., 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

[4] Chen, Y.H. 2005. *The Pitch Contours of Infant Babbling Adapting to the Target Language*. MA thesis. Linguistics Department, Hunan University.

[5] Clumeck, H., 1980. The acquisition of tone. In *Child Phonology. Volume 1:Production*. G.Yeni-Komshian, J. Kavanagh and C. Ferguson (eds.). London: Academic Press.

[6] Ferguson, C.A.; Farwell, C.B., 1975. Words and sounds in early language acquisition. *Language* 51(2), 419-439.

[7] Hsieh, H.I., 1977. Lexical diffusion: evidence from child language acquisition. In *The Lexicon in Phonological Change*, W.S-Y.Wang (ed.). The Hague: Monton Publishers, 133-148.

[8] Hsu, J., 1996. A study of the stages in the development and acquisition of Mandarin Chinese by children in Taiwan. Taipei: The Craine Publishing, Co., Ltd.

[9] Lee, T.; Lee, B.; Chen, J., 2005. The pitch patterns of infants acquiring tonal languages: the perspective from Chinese. Paper presented at ELA2005, University of Lyon, France, Dec 8-10, 2005.

[10] Li, C.; Thompson, S., 1977. The acquisition of tone in Mandarin-speaking Children. *Journal of Child Language* 4,185-199.

[11] Stoel-Gammon, C. Cooper, J., 1984. Patterns of early lexical and phonological development. *Journal of Child Language* (11), 247-271.

[12] Tsay, J., 2004. The production of Taiwanese tone sandhi in child language. Paper presented at the Language Acquisition Workshop, Graduate Institute of Linguistics, Tsinghua University, March 16, 2004.

[13] Tse, J. K.-P., 1977. Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language* (5), 191-204.

[14] Tse, A. C.-Y., 1993. The development of a phonological System in Cantonese: A case report. In *The proceedings of the 25th Annual Child Languae Research Forum*, Eve Clark (ed.). Stanford, CA, Center for the Study of Language and Information.

[15] Vihman, M. M. (1996). *Phonological development: the origin of language in the child*. Oxford, Blackwell.

[16] Wong, P.-S., Schwartz, R.G., Jenkins, J..J. 2005 Perception and production of lexical tones by 3-year-old Mandarin-speaking children. *JSLHR*, vol. 48, 1065-1079.

[17] Woo, N., 1969. Prosody and phonology. Doctoral dissertation, MIT, Cambridge, Mass.

[18] Yip, M., 1980. The tonal phonology of Chinese. Doctoral dissertation, MIT, Cambridge, Mass.

[19] Yip, M., 1989. Contour tones. *Phonology*, 6.1:149-74.

[20] Zhu, H.; Dodd, B., 2000. The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language* 27, 3-42.

[21] Zhu, H., 2002. *Phonological development in specific context: studies of Chinese-speaking children*. Clevedon: Multilingual Matters Ltd.

[22] Zhu, X.N. 2004. Jipin Guiyihua (Normalization of Fundamental Frequency). *Yuyan Kexue (Language Sciences)* 9, 3-19.

* We are indebted to various members of the Hunan Chinese Early Language Acquisition group who contributed to the recording of AJR: Ai Zhaoyang, Zeng Tao, Huang Aijun, Cai Xin, Chen Feiyan, Chen Min, Chen Yanhua, Liao Hui, Xiao Ling and Yang Jie. We owe special thanks to Ning Chunyan for his staunch support of acquisition research at Hunan University. We are grateful to Wang Hao and Jane Tsay for technical advice, and to Ying-wai Wong and two anonymous reviewers for comments on an earlier version of this paper. The research reported in this paper was supported by two grants to Thomas Lee, an RGC grant (CityU 1245/02H), and a grant from the Arts Faculty of the Chinese University of Hong Kong (CUHK 2010282).