A preliminary study on the acquisition of Mandarin neutral tone by young heritage children

Chen Lan, Peggy Mok

The Chinese University of Hong Kong

`lchsapphire@gmail.com, peggymok@cuhk.edu.hk`

Abstract

The present study examined the pitch and duration of three types of Mandarin neutral tone (T0) – the possessive particle -de, the noun suffix -zi, and reduplicated words, in four tonal environments (following T1/T2/T3/T4) produced by two heritage language (HL) children longitudinally at 3;0 and 4;0 and two children cross-sectionally at 4;0, comparing them with two adult native speakers. Unlike the monolingual children in a previous study, our results indicated that the HL children have not developed a robust neutral tone category by 4;0. Acoustic measurements showed that HL children shared similar shapes of neutral tone pitch contour across types and tonal environments with the adult speakers when they were at 4;0, with a falling pitch contour following T1, T2, T4 and a rising pitch contour following T3. However, even though both HL children and the reference speakers reduced duration while producing neutral tone compared to the preceding tone, significant differences were found for their duration variations; while the reference speakers produced neutral tone after T1, T2, and T4 with a shorter duration than T3 across types, the patterns of HL children varied according to different types of the neutral tone. Individual difference also existed among the HL children.

Index Terms: neutral tone, heritage language, Mandarin Chinese

1. Introduction

Heritage language (HL) is a newly recognized form of bilingualism, which has attracted growing research interest in recent years. Heritage speakers (HSs) hear and speak their HL as well as the majority language sequentially or simultaneously in early childhood. Several studies investigated HL phonology in Mandarin Chinese, including both vowels and tones, but they primarily focused on teenagers or adult speakers, e.g. [1, 2]. Research on young children, in particular those at the early stage of language development, is still lacking. Longitudinal and naturalistic corpus data provide a valuable pool for tracking the language development across ages, contributing to better understanding of how sound system develop by young heritage children. This study made use of the corpus data to investigate the acoustic characteristics of neutral tone by young heritage children at 3;0 and 4;0, aiming to establish a preliminary understanding on the acquisition of neutral tone system by heritage children.

1.1. Acquisition of Mandarin neutral tone

Mandarin Chinese has four lexical tones contrasting in pitch contours: T1 [55] a high level tone, T2 [35] a rising tone, T3 [214] a dipping tone, and T4 [51] a falling tone [8]. In addition to these four tones, there is a T0 neutral tone occurring on weak syllables or the final position of a word, with a short duration and a reduced pitch range, being influenced by the preceding tone [3, 4, 5, 6]. Neutral tone can be divided into three types when appearing in different semantic contexts, including the reduplication type in which the second syllable of the reduplicated word carries T0, the suffix type (i.e. the possessive particle -de), and the lexeme type in which the second lexeme of a two-lexeme word carries T0 [4, 5, 6].

Previous studies on native Mandarin-speaking children suggested that there are four categories of neutral tone appearing in the child's vocabulary, including the classifier -ge, the particle -de, the reduplicated words, and the noun suffix -zi [4]. With regard to the acquisition of T0, some researchers proposed that monolingual Mandarin-speaking children were able to produce the neutral tone as early as 1;2, but the acquisition is not fully completed at 4;6 [6, 7]. Several error patterns in child production were reported; for example, monolingual children usually used citation tone in the place of the neutral tone, lengthened the first syllable of a disyllabic word, or deleted the syllable that carry the neutral tone [4, 6, 7]. The other researchers argued that even though monolingual children do not have adult-like productions at 5;0, they have developed a phonological category of neutral tone by 3;0 [5]. For instance, in adults' productions, T0 had a falling pitch contour after T1, T2, T4 and a level/rising pitch contour after T3; T0 was produced with a longer duration after T3 than after T1, T2, and T4. The productions of monolingual children at 3;0 shared a similar pitch and duration pattern as the adult speakers', and they became closer to the adult pattern as they got older [5]. In terms of the acquisition of heritage Mandarin, only one published study examined the production of neutral tone by adult heritage speakers using acoustic analysis and compared with L1 speakers and L2 leaners of Mandarin, revealing little difference in pitch contour but significant differences in duration [8]. Also, HSs were not more adult-like compared to the L2 learners, and their performances depended on whether a token was in obligatory or non-obligatory T0 context. However, how young heritage children acquire the neutral tone remains unknown. This study will fill this gap by investigating the production of neutral tone by young heritage children.

1.2. Phonological development of heritage language

There are several issues relevant to phonological studies on heritage language, and they were also addressed in this study. First, among the current studies on heritage Chinese phonological development, most of them focused on young adults or teenagers whose HLs were no longer developing, investigating whether HSs maintain the sound system as the native speakers or converge to the sound system of the majority language of the society after the onset of schooling when they are systematically exposed to the dominant language. However, considering that most children may attend daycare or kindergarten, language shifting may appear earlier at preschooler age [9]. Therefore, in addition to the ultimate attainment of HSs, research on the early stages of language

development in childhood is also important for a more comprehensive understanding of HL development. Study on child speech acquisition of heritage language is warranted.

The second issue of HL phonological studies is related to the phonological advantages. Some researchers suggested that heritage speakers have an advantage over the late L2 learners of the target language on both production and perception since they have a similar acquisition method as the monolingual speakers, acquiring the target language since birth at home (e.g. [10, 11]). However, some other studies indicated that HSs may not have an advantage over the learners of the target language, and whether they have such an advantage highly depend on the HL experience of HSs [8]. With mixed findings, more studies on HL phonological development are needed.

Moreover, a latest issue regarding the lifespan development of HSs' phonetic and phonological system calls for the longitudinal research tracking HSs over time [12], encouraging the researchers to use longitudinal data for HL phonological studies. This study took a preliminary step using longitudinal data to investigate the acquisition of heritage Mandarin.

To sum up, the present study examined the Mandarin neutral tone production by two heritage children longitudinally at 3;0 and 4;0 and two HL children cross-sectionally at 4;0, aiming to conclude their neutral tone patterns, compare them with those of native Mandarin speakers, and track their development across age. Both linguistic and social factors affecting the attainment of HSs were considered.

2. Method

2.1. Participants

Data for this study came from the Bilingual Child Heritage Chinese Corpus (CHCC; [13]) available in the CHILDES corpus(https://childes.talkbank.org/access/Biling/CHCC.html). CHCC has released longitudinal data recorded from three heritage speakers of Chinese from age 1;07 up to 4;11. All of them were born in United States, being exposed to Mandarin since birth at home. Considering the influence of Cantonese inputs on the acquisition of neutral tone, the child named Winston with Cantonese background was excluded. Avia and Luna's data were analyzed. In addition to the released data, more data from two children - Fiona and Nathan, collected with the same method as the released data, were transcribed and analyzed in this study. In total, two HL children longitudinally at 3;0 and 4;0 and two HL children cross-sectionally at 4;0 were included. Table 1 summarizes the basic information of the HL children. According to [13], Luna, Fiona, and Nathan shared similar language environment, growing up with both parents speaking Mandarin at home. Avia was raised in a family adopting "one parent-one language" principle with native Mandarin spoken by her mother and native English spoken by her father; thus, she received native Mandarin and English inputs since birth at home. Considering that the adult patterns of T0 have been well described in the literature, only two native adult speakers of northern Mandarin (mean age = 26.5) were included as the reference speakers in this study.

Table 1: *Information of the HL participants.*

Subject	Gender	Age range of data used
Avia (released)	F	2;11 – 3;2
		3;10 – 3;11
Luna (released)	F	2;11 – 3;2
		3;11 – 4;3
Fiona	F	4;2 – 4;3
Nathan	M	3;10 – 4;0

2.2. Procedure

For Avia and Luna, the online transcripts of the recordings were used. Disyllabic words ended with the possessive particle *-de* and the noun suffix *-zi*, and the reduplicated words which potentially carrying the neutral tone produced by each child were listed (exemplified in Table 2). For Fiona and Nathan, simple transcriptions were done by the author to access the target disyllabic words. Each sound file containing the target types of T0 and being judged by the author with phonetic training as an acceptable neutral tone production were annotated and processed. Considering that the acoustic analysis would be sensitive to sound quality, the files with heavy background noise were excluded. In total, 76 disyllabic words produced by two heritage children at 3;0, consisting of 41 tokens of suffix type and 35 tokens of reduplication type, and 100 disyllabic words produced by four heritage children at 4;0, including 75 tokens of suffix type and 25 tokens of reduplication type, were analyzed. The data were also divided into four tonal environments, consisting of 49 T1+T0 words, 42 T2+T0 words, 40 T3+T0 words, and 45 T4+T0 words.

Table 2: *Target types of the neutral tone*

Type	Function	Example	Citation tone
Suffix	Normalizer / possessive particle *-de*	wo3*de0* 'mine' gou3*de0* 'dog's'	
	Noun suffix *-zi*	zhuo1*zi0* 'table'	/tsi3/
Reduplication	The second syllable of the reduplicative word.	ma1*ma0* 'mother' xing1*xing0* 'star'	/ma1/

The data from the reference speakers were collected with a reading task. The stimuli were selected from the dissyllabic words produced by the children. The speakers were instructed to read out the stimuli isolated twice. The recording was conducted in a quiet room. In total, 115 tokens, containing the target types and covering four tonal environments, were analyzed and compared with the heritage data.

2.3. Data analysis

This study analyzed two prosodic parameters of neutral tone – pitch and duration. The acoustic measurements were done using ProsodyPro [14] in Praat [15]. Fundamental frequency (F0) was measured using the automatic F0 tracking on 10 equidistant points along the tone contour. All data were measured in semitones. Manual pulse fixing was done to ensure data validity. Considering that some of the tokens produced by the participants contained the glide, such as /j/ in /je2 *je0*/ "grandfather", instead of comparing the raw duration of T0, this study followed [5] evaluating the normalized duration, which was a ratio between the duration of the neutral tone and the duration of the preceding tone. All data were measured in milliseconds (ms).

The statistical analysis was done with linear mixed effect model using lme4 package [16] in R [17]. The model used Group (3-year-old heritage speakers, 4-year-old heritage speakers, and the reference speakers), Type (suffixation and reduplication), and tonal Environment (T1+T0, T2+T0, T3+T0, T4+T0) as fixed effects and Speaker as the random effect. Separate pairwise t tests (with Bonferroni correction) were for further comparisons.

3. Results

3.1. Pitch

Figure 1 shows the average pitch contours of the preceding lexical tones (T1, T2, T3, T4) and the following neutral tone (T0) produced by 3-year-old heritage children, 4-year-old heritage children, and the reference speakers. Overall, all groups produced T0 with a falling pitch contour after T1, T2, and T4 and with a rising contour after T3. However, the pitch contours of T0 produced by 4-year-old heritage children are more neutralized compared to that of 3-year-old heritage children as well as the reference speakers. The distance between T0 contours after T1 and T2 by heritage children was observed to be closer than that of reference speakers.

Further evaluations focused on T0 according to two types of neutral tone – Suffixation (the possessive particle -de and the noun suffix -zi) and Reduplication. As shown in Figure 2, for both suffixation type and reduplication type of neutral tone, T0 was produced by HL children with a falling pitch contour after T1, T2, and T4 and with a rising pitch contour after T3, which is similar to the reference pattern. However, for suffixation type produced by 4-year-old HL children, the first half of the T0 contours following T1 and T2 were observably overlapping, and the T0 pitch slope in T1+T0 tonal environment was not as steep as the slope by 3-year-old children and the reference speakers.
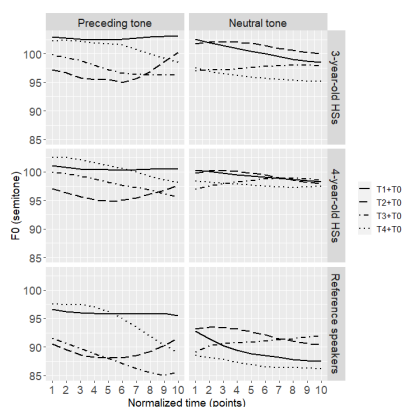

Figure 1: *Mean pitch contour of the preceding lexical tones (left) and the neutral tone (right) in four tonal environments by 3-year-old heritage children (up), 4-year-old heritage children (middle), and the reference speakers (bottom).*
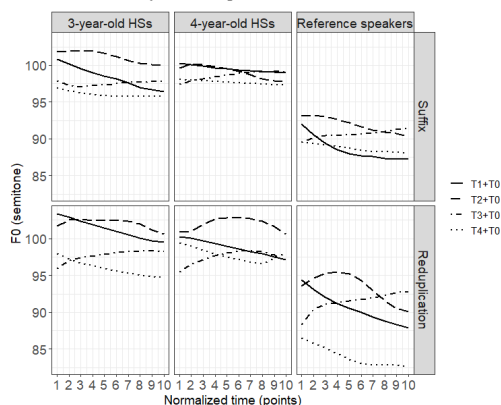

Figure 2: *Mean pitch contour of the neutral tone in four tonal environments across two types (up: Suffix; bottom: Reduplication) by 3-year-old heritage children (left), 4-year-old heritage children (middle), and the reference speakers (right).*

Table 3: *Normalized duration (in ms) of two neutral tone types produced by 3-year-old heritage children, 4-year-old heritage children, and the reference speakers.*

	Suffix	Reduplication	*p* value
3-year-old HSs	0.57 (0.04)	0.81 (0.05)	< .001
4-year-old HSs	0.56 (0.03)	0.93 (0.08)	< .001
Reference speakers	0.48 (0.02)	0.62 (0.03)	< .001

3.2. Duration

Figure 3 indicates the normalized neutral tone duration in four tonal environments produced by 3-year-old heritage children, 4-year-old heritage children, and the reference speakers. Generally, both groups of heritage children as well as the reference speakers reduced duration while producing neutral tone compared to the preceding tone (ratio < 1). Linear mixed effect model on the normalized duration, using Group, Type, and tonal Environment as fixed effects and Speaker as random effect, revealed a significant difference among groups ($|t|$ = 2.282, p = 0.03) and among tonal environments ($|t|$=1.973, p = 0.05). For speaker groups, there was a significant difference between reference speakers and 3-year-old HL children (p < .001) as well as between reference speakers and 4-year-old HL children (p = .001). No significant difference was found between the two groups of HL children. For different tonal environments, the reference speakers produced T0 after T3 with a relatively longer duration, followed by T1, T2, and T4. For the HL children, T0 produced by 3-year-old children was relatively longer after T3, followed by T1, T4, and T2; T0 produced by 4-year-old children was relatively longer after T1, followed by T3, T4, and T2, differing from the reference pattern. Separate pairwise t tests further showed that the reference speakers produced T0 following T3 with a significantly longer duration than following T1 (*p* = .050), T2 (*p* < .001), and T4 (*p* < .001). The duration of T0 after T4 was also significantly shorter than after T1 (*p* = .028). Unlike the reference speakers, even though 3-year-old HL children produced T0 after T3 with a numerically longer duration than after T1, T2, and T4, significant difference was only found between T3 and T2 (*p* = .036). T0 in T4+T0 environment produced by 4-year-old heritage children was numerically longer than in other tonal environments; however, no significant difference was found between each two of the environments.

Further examinations were done on different types of neutral tone. Table 3 shows the normalized T0 duration of each type produced by three groups of speakers. Separate pairwise comparisons revealed significant differences between two types of neutral tone by both heritage children and the reference speakers (*p* < .001), with a significantly shorter duration for the suffix type than for the reduplication type. For the suffix type, a significant main effect was found for Group (F (2,192) = 3.358, *p* = .037), but there was no significant difference between each two groups. For the reduplication type, there was a significant main effect of Group (F (2, 93) = 9.373, *p* < .001). T0 produced by the reference speakers was significantly shorter than that of 3-year-old HL children (*p* = .018) and 4-year-old HL children (*p* < .001). For the reference speakers and 3-year-old HL children, T0 after T3 was produced with a relatively longer duration than after T2, T1, and T4 across types, while 4-year-old heritage produced reduplication type of T0 after T2 and suffix type after T1 with a longer duration (see Figure 4). More specifically, for the reference speakers, the normalized

duration of the suffix neutral tone following T3 was significantly longer than T2 ($p = .011$) and T4 ($p = .003$); the suffix type following T4 had a significantly shorter duration than T1 ($p = .035$). The normalized duration of reduplicative T0 after T3 was significantly longer than after T1 ($p = .003$) and T4 ($p < .001$). However, for both types, no significant difference on normalized T0 duration was found between each two of the tonal environments by heritage children.
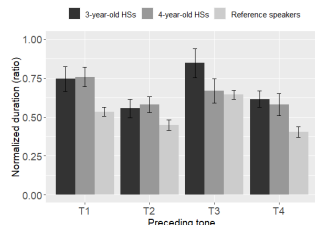

Figure 3: *Average normalized duration of T0 in four tonal environments produced by 3-year-old heritage children, 4-year-old heritage children, and the reference speakers.*
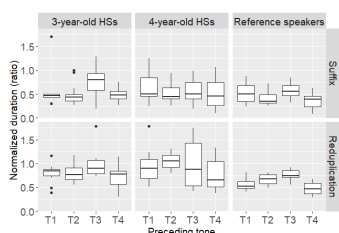

Figure 4: *Normalized duration of suffix type (up) and reduplication type (bottom) in four tonal environments produced by 3-year-old heritage children (left), 4-year-old heritage children (middle), and the reference speakers (right).*
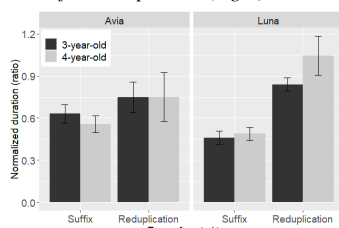

Figure 5: *Average normalized duration of suffix type and reduplication type of T0 produced by Avia (left) and Luna (right) at 3;0 and 4;0.*

3.3. Longitudinal development

Different patterns were found for Avia and Luna who were longitudinally tracked for one year (see Figure 5). Avia shortened the syllable carrying suffix type of T0 as age increased, becoming closer to the reference pattern. In contrast, compared to the productions collected at 3;0, Luna produced both suffix type and reduplication type of neutral tone with a longer duration at 4;0. However, no significant difference was found between two age ranges. Also, for most of the reduplicated words produced by Luna at 4;0, in particular in the T2+T0 and T3+T0 tonal environments, the ratio between T0 duration and the preceding lexical tone duration was larger than 1, indicating that she did not reduce duration when producing T0 compared to T2 and T3 in reduplicative words, which differed from the reference pattern.

4. Discussion

This study evaluated the production of Mandarin neutral tone by 3-year-old and 4-year-old HL children, reporting a generally similar contour pattern but a different duration pattern compared to the native Mandarin reference speakers. The 3-year-old and 4-year-old HL children produced T0 with a falling pitch contour after T1, T2, T4 and a rising contour after T3, which is consistent with the pattern of the monolingual Mandarin-speaking children with comparable ages in [5]. However, while the reference speakers and the monolingual children in [5] produced significantly longer T0 after T3 than other tones, the duration patterns of HL children, especially the 4-year-olds are different. This finding concurs with the finding in terms of the comparison among adult HSs and L1 speakers in [8], with little difference in pitch contour but significant difference in duration. It is possible for heritage children to produce and perceive neutral tone relying more on pitch than duration, but further study is needed to confirm this point.

For both pitch contour and duration, compared to the 4-year-old pattern, the 3-year-old pattern was observed to be more similar to the reference pattern. For example, the pitch contours of 4-year-olds were more neutralized than that of other groups. The duration pattern of 3-year-olds was similar to the reference pattern across types, with T0 after T3 was longer, while the 4-year-old pattern was different, with T0 after T1 or T2 was longer. Also, the reduplication type produced by 3-year-old children was shorter than that of 4-year-olds. These findings on HL children suggested that the phonological development of neutral tone was not as consistent and continuous as the monolingual children in [5] that children became closer to the adult speakers as age increased. There are some possible explanations. First, as the children get older, there might be less opportunity for them to perceive and produce reduplicated words. Among the corpus data used in this study, there were 35 tokens of reduplication type in the recordings for 3-year-old group, but only 25 tokens could be extracted for 4-year-old group. The decrease in the HL experience on the reduplication type may influence the accuracy and the developmental pattern of their productions. The second possible factor may be the influence of the majority language, which is English in the United States. Take Luna's case for example - starting from age 3;02, Luna was attending an English-medium pre-school, which might lead to an increase in the English inputs. The main language input for her might change from Mandarin to English or to English-Mandarin, causing an interaction of the Mandarin and English prosodic systems, which influence her productions. In addition, a stable development of neutral tone may be affected by whether the children receive pure neutral tone inputs. For instance, compared to Avia whose T0 inputs came from her mother who speaks northern Mandarin, Luna's mother speaks southern Mandarin which may not have neutral tone, causing a decrease in her T0 inputs and influencing her development. Also, when the HL children gradually joined more social activities as they became older, they may meet other Chinese who speak Mandarin without neutral tone, which may also affect their productions.

In conclusion, this study investigated the production of Mandarin neutral tone produced by two HL children longitudinally at 3;0 and 4;0 and two HL children cross-sectionally at 4;0 in the United States. HL children shared similar overall pitch patterns with the reference speakers and the monolingual children in [5], but with different duration patterns. More data will be examined to corroborate the findings discussed in this preliminary study.

5. References

[1] C. B. Chang, Y. Yao, E. F. Haynes, and R. Rhodes, Production of phonetic and phonological contrast by heritage speakers of Mandarin. *The Journal of the Acoustic Society of Ameirca, 129*(6), 3964 – 3980, 2011.

[2] C. B. Chang and Y. Yao, Toward an understanding of heritage prosody: Acoustic and perceptual properties of tone produced by heritage, native, and second language speakers of Mandarin. *Heritage Language Journal, 13*(2), 134-160, 2016.

[3] A. Li, Phonetic correlates of neutral tone in different information structures. *Contemp. Linguist. 19*(3), 348-378, 2017.

[4] C. N. Li and S. A. Thompson, The acquisition of tone in Mandarin-speaking children. *Journal of Child Language, 4*(2), 185–99, 1977.

[5] P. Tang, I. Yuen, N. Xu Rattanasone, L. Gao, and K. Demuth, Acquisition of weak syllables in tonal languages: acoustic evidence from neutral tone in Mandarin Chinese. *Journal of Child Language* (2018), 1–27, 2018.

[6] Z. Hua and B. Dodd, The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language, 27*(1), 3–42, 2000.

[7] Z. Hua, *Phonological development in specific contexts: studies of Chinese-speaking children*. Clevedon: Multilingual Matters, 2002.

[8] C. B. Chang and Y. Yao, Production of neutral tone in Mandarin by heritage, native, and second language speakers. In *Proceedings of the 19th International Congress of Phonetic Sciences ICPhS 2019*. Melbourne. 2019.

[9] A. Stoehr, T. Benders, J. G. Van Hell, and P. Fikkert, Heritage language exposure impacts voice onset time of Dutch German simultaneous bilingual preschoolers. *Bilingualism: Language and Cognition, 21*(3), 598-617, 2017.

[10] J. S. Oh, S. A. Jun, L. M. Knightly, and T. K. Au, Holding on to childhood language memory. *Cognition, 86*(3), B53-64, 2003.

[11] L. M. Knightly, S. A. Jun, J. S. Oh, and T. K. Au, Production benefits of childhood overhearing. *JASA 114*(1), 465–474, 2003.

[12] C. B. Chang, Phonetics and phonology. In S. Montrul & M. Polinsky (Eds.), *The Cambridge Handbook of Heritage Languages and Linguistics* (Chapter 23). Cambridge, UK: Cambridge University Press. 2021.

[13] Z. Mai and V. Yip, Acquiring Chinese as a Heritage Language in English-speaking Countries and the Child Heritage Chinese Corpus. Paper presented at *the International Conference on Bilingualism: Language and Heritage*, Dec 18, Chinese University of Hong Kong. 2017.

[14] Y. Xu, ProsodyPro- A Tool for Large-scale Systematic Prosody Analysis. In Proceedings of Tools and Resources for the Analysis of Speech Prosody, TRASP, 2013.

[15] P. Boersma and D. Weenink, Praat: doing phonetics by computer [Computer program]. Version 6.2.01, retrieved from http://www.praat.org/, 2021.

[16] D. Bates, M. Mächler, B. Bolker, and S. Walker, Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48, 2015.

[17] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. 2021.