J. Child Lang. **37** (2010), 341–371.
doi:10.1017/S0305000909009581

Segmental production in Mandarin-learning infants*

LI-MEI CHEN

National Cheng Kung University, Taiwan, ROC

AND

RAYMOND D. KENT

University of Wisconsin-Madison

(*Received* 6 *December* 2004 – *Revised* 16 *November* 2008 – *Accepted* 19 *March* 2009 – *First published online* 3 *June* 2009)

ABSTRACT

The early development of vocalic and consonantal production in Mandarin-learning infants was studied at the transition from babbling to producing first words. Spontaneous vocalizations were recorded for 24 infants grouped by age: G1 (0;7 to 1;0) and G2 (1;1 to 1;6). Additionally, the infant-directed speech of 24 caregivers was recorded during natural infant–adult interactions to infer language-specific effects. Data were phonetically transcribed according to broad categories of vowels and consonants. Vocalic development, in comparison with reports for children of other linguistic environments, exhibited two universal patterns: the prominence of [ɛ] and [ə], and the predominance of low and mid vowels over high vowels. Language-specific patterns were also found, e.g. the early appearance and acquisition of low vowels [ɑ]. Vowel production was similar in G1 and G2, and a continuum of developmental changes brought infants' vocalization closer to the adult model. Consonantal development showed two universal patterns: labials and alveolars occurred more frequently than velars; and nasals developed earlier than fricatives, affricates and liquids. We also found two language-specific patterns: alveolars were more prominent than labials and affricates developed early. Universal and language-specific characteristics in G1 continued to be prominent in G2. These data indicate that infants are sensitive to the ambient language at an early age, and this sensitivity influences the nature of their vocalizations.

[*] We are grateful to the infants and their families who participated in this study. Without their cooperation, this study would not have been possible. Their hospitality made the whole recording procedure an enjoyable experience. Address for correspondence: Raymond D. Kent, Waisman Center, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, WI, 53705, USA. e-mail: Kent@Waisman.Wisc.edu

INTRODUCTION

Early phonetic behavior appears to be determined by three primary factors: (1) articulatory competence, related especially to developing vocal tract anatomy and motor function; (2) influence of the ambient language, mediated by auditory perceptual processes; and (3) neural capability, both as it relates to the first two factors and as it pertains to other aspects of psychobiological development underlying communication. Since each factor has a developmental course, typical or normal sound production in infancy reflects the balance and interaction among articulatory proficiency, linguistic influence and neural status. Furthermore, because these factors are closely interwoven, it is difficult to identify their separate contributions, except through cross-linguistic studies or studies of clinical populations (e.g. infants with craniofacial anomalies, hearing impairment or cognitive disability). A developmental period of particular interest is the transition from babbling to first words, when the influence of ambient language can be gauged with some confidence, and when an infant's vocal tract, as well as motor control of vocal tract structures, takes on some important properties of the adult system.

A fundamental phonetic divide is that between vowels and consonants. The vowel–consonant distinction is manifold, including differences in auditory processing, articulatory motor activity, relative age of acquisition and susceptibility to developmental speech sound errors. Specifically, compared to consonants, vowels are: (1) processed perceptually in different cortical regions (Carreiras & Price, 2008); (2) acquired earlier in development (Stoel-Gammon & Pollock, 2008); and (3) less vulnerable to articulation errors (Chen & Irwin, 1946). Despite these differences, vowels and consonants appear together in speech development, especially in CV syllables, which have been suggested to be an important phonetic training ground for speech development. Furthermore, the appearance of both vowels and consonants in infant vocalizations has been described in terms of strong biological constraints resulting in a near universality of some sounds in babbling and the virtual exclusion of others (Locke, 1983). Eventually, the vowels and consonants in infant vocalizations, particularly in early words, reflect the phonetic properties of the ambient language. An infant's audition mediates between the sounds of adult language and the infant's repertoire of sound production. During the period of first words, if not before, infants demonstrate sensitivity to vowel and consonant elements in tasks of word detection (Mani & Plunkett, 2008).

The promise of cross-linguistic studies to elucidate these issues has been only modestly realized. In particular, limited data have been published on the acquisition of speech sounds in Asian languages. Mandarin is of interest in comparative developmental linguistics because it exhibits overwhelmingly monosyllabic structures along with a high frequency of low vowels and

https://doi.org/10.1017/S0305000909009581 Published online by Cambridge University Press

dental consonants (Cheng, 1982), features that contrast with the high frequency of front vowels and labial consonants in English (Vihman, Kay, de Boysson-Bardies, Durand & Sundberg, 1994). CV syllables are of core interest because they are often regarded as a critical training ground for speech development and may be core units in early lexical acquisition.

This study examines the acquisition of vowels and consonants in early CV syllables produced by Mandarin-learning infants. The findings of this study are relevant to long-standing developmental questions: universal vs. language-specific characteristics and continuity vs. discontinuity in phonetic development. The universal characteristics are identified by comparison with results from other language groups. The language-specific effects are inferred from comparison with the major linguistic input, child-directed speech.

Universal vs. language-specific development

Common characteristics of vowels. This section reviews universal and language-specific characteristics in vowel production from cross-linguistic studies and those on the acquisition of Mandarin vowels. In particular, major findings are discussed on vocalic development in the first year of life and from the first year to the appearance of early meaningful speech. Due to similar courses of development in articulatory proficiency for children learning English and other languages, their data often show common characteristics in vowel development. Indeed, early vowels have been shown to be mainly central and front vowels. For example, in an early study of vowel development based on transcriptions of the vocalizations from 95 English-learning infants (Irwin, 1948), mid-front [ɛ, ɪ] and mid-central (ə) vowels accounted for over 70% of vowel production during the first year. Among these vowels, [ɛ] occurred most often. The major trend in vowel development, according to Irwin, is an increasing production of back vowels from age 0;11 to 2;6.

Similarly, analysis of vocalizations by five infants aged 1;1 showed that central and front vowels [ʌ (ə), ɛ, æ] are produced most often in V (monophthongs), VV (diphthongs) and CV (vowels with an initial consonant) syllables (Kent & Bauer, 1985). In addition to these three vowels, [ɑ] is often produced in CV syllables. These authors concluded that central and front vowels are produced more often than back vowels, and low vowels are produced more often than high vowels.

This developmental trend has been supported by most studies analyzing formant patterns. For example, Buhr's (1980) longitudinal study found that the vowel spaces (defined by F1 and F2 values) of non-high front and central vowels [ɪ, ɛ, æ, ʌ] become stabilized by the end of the first year, earlier than the stabilization of any back or high vowels. Furthermore, Kent

https://doi.org/10.1017/S0305000909009581 Published online by Cambridge University Press

& Murray's (1982) analysis of 21 infants aged 0;3, 0;6 and 0;9 indicated that the vowel formant frequencies of younger children are more upward and to the right in the standard F1–F2 graph than those of older children and adults. That is, vowels produced from ages 0;3 to 0;9 are mainly mid-front or central vowels. This trend is supported by Matyear's (1998) study of vowel formants in three English-learning infants (ages 0;7 to 1;0) who produced mostly non-high, non-back vowels.

Similar results are found in Mandarin-learning infants. Mid-front and mid-central vowels [ɛ, ə], often used in babbling before age 0;8 (Chen, 2005), have also been reported in first words by age 1;8 (e.g. Jeng, 1979; Yue-Hashimoto, 1980). The emergence of the contrast of three corner vowels [i, ɑ, u] before age 2;0 (e.g. Jeng, 1979; Hsu, 2003) must be the major development during the second year, since few utterances of both [i] and [u] have been observed before age 1;0 (Chen, 2005).

Language-specific characteristics of vowels. In addition to these common characteristics, language-specific vowel development is found in cross-linguistic studies and in studies of Mandarin-learning children.

Language-specific vowel production occurs by age 0;10 (de Boysson-Bardies, Hallé, Sagart & Durand, 1989; de Boysson-Bardies, 1993). The F2/F1 ratio of vowel production of 20 infants (aged 0;10) from English, French, Algerian and Cantonese environments reflected the F2/F1 ratio in their corresponding adults' models (de Boysson-Bardies *et al.*, 1989). For example, among these four language groups, the vowel production of both English-learning infants and their corresponding adult language group exhibited the highest F2/F1 ratio, reflecting the high frequency of mid- to high-front vowels [ɪ, e] in the adult model. In contrast, the vowel production of infants in the Cantonese environment showed the lowest F2/F1 ratio, correlating with the lowest F2/F1 ratio of their adult model and reflecting the high frequency of low-back vowels [ɑː, ɔː, ɐ]. Similarly, in Stokes & Wong's study (2002) of 40 Cantonese-speaking children from age 0;10 to 2;3, the three vowels [a], [ɛ] and [ɔ] were the first to emerge, achieving better than 50% accuracy by age 1;3 to 1;6. These were followed by [i] at age 1;8 to 1;11.

In addition to spectral characteristics and the distribution of vowels, language-specific influences in infants' vowel production have been examined in terms of height and place relations of V1 and V2 in CVCV sequences (de Boysson-Bardies, 1993). In this second cross-linguistic study, the V1 and V2 elements in disyllabic utterances of infants and adults from four language environments (French, English, Swedish and Yoruban) tended to agree in height and frontedness. Utterances with vowels that vary with height from French-learning and English-learning infants showed a tendency toward higher V2, which is related to their adult models. Disyllabic productions of infants from Swedish and Yoruban environments

TABLE 1. *Single vowels in Mandarin*

	Vowel backness				
Vowel height	Front		Central	Back	
High	i	y	ɨ (2 variants)		u
Mid	e		ə	ɤ	o
	(ɛ)				(ɔ)
Low			a	ɑ	

Lin's (1989) analysis of Mandarin surface vowels.
([ɛ] and [ɔ] are only used in forming diphthongs).

displayed a tendency for a higher V1 to occur at high frequency, reflecting their corresponding adult models. However, no such systematic correlation was found in the variation of frontedness of V1 and V2. In conclusion, language-specific vowel production begins in babbling by age 0;10, shortly before first words are produced.

Several language-specific characteristics in Mandarin-speaking children have also been reported but before these studies are reviewed, the Mandarin vowel system is introduced. Although analyses of the Mandarin vowel system remain unsettled on the number of surface vowels and distribution of allophones in the phonemic system, most studies propose 12 or 13 surface vowels and 4 to 6 vowel phonemes (Cheng, 1966; Cheng, 1973; Lin, 1989). These analyses differ primarily with respect to the low vowels /a/ and /ɑ/ and the high vowel /ɨ/. All analyses regard [a] and [ɑ] as variants of the same vowel phoneme, the only low vowel in Mandarin. The low-back vowel /ɑ/ is considered a basic form by Cheng (1973), whereas others propose /a/ as a phoneme, with [a, ɑ, ɛ] or [æ] as allophonic variations. The high vowel /ɨ/ has been regarded as phonemic and having two variants (Cheng, 1966; Cheng, 1973), or as an allophone of /i/. All 12 surface vowels of Cheng (1966) and Cheng (1973) (i, e, ɛ, y, ɨ [with 2 variants], ə, a, u, ɤ, o, ɑ) and an extra [ɔ] are included in Lin's (1989) analysis of Mandarin surface vowels (Table 1). The same set of vowels occurs in Mandarin spoken in Taiwan, except that the two allophonic variations of the high vowel /ɨ/ are not distinguished.

To analyze the frequency of Mandarin syllables, this study used a corpus of 1 177 984 Mandarin characters from written materials (e.g. journals, newspapers, textbooks; Liu, Chuang & Wang, 1975; Cheng, 1982), since no spoken corpus is available. Our analysis of this database, including over 900 000 syllables with both consonant and vowel components, shows that front vowels occur most frequently (44·65%), followed by central (31·48%) (including [a], if the low-central vowel [a] is considered the basic form) and back vowels (23·86%). As to vowel height, high vowels occur most frequently (42·09%), followed by mid vowels (35·21%) and low vowels (22·69%). The low-central vowel [a] (22·69%) and the high-front vowels [i, y] (25·80%) occur most frequently among all vowels.

Chen's (2005) finding that [ɛ] is produced early and that high and back vowels [i, ɪ, u, ʊ, o, ɔ] are rarely produced (less than 7%) has also been documented in English-learning infants (Irwin, 1948; Buhr, 1980; Kent & Bauer, 1985). The early emergence and increasing production of the low vowel [a] or [ɑ] by a Mandarin-learning infant (0;3–0;8) reflects the phonetic system of the ambient language (Chen, 2005). This increasing production of [ɑ] is like that of Cantonese-learning infants (age 0;10) (de Boysson-Bardies *et al.*, 1989). Moreover, [a] or [ɑ] is already present in Mandarin infants' vowel inventories by the second year (e.g. Li, 1978; Yue-Hashimoto, 1980; Wang, Huang, Chen & Phillips, 1986).

Cross-linguistic differences in vowel production in babbling are found as early as age 0;10, several months before the first words are produced. For example, in addition to the general trend for mid-front and mid-central vowels to occur with high frequency in Mandarin-learning infants, production of [a] or [ɑ] increases before age 1;0, corresponding to the prominence of low vowels in Mandarin syllables mentioned above.

Common characteristics of consonants. Common developmental patterns of consonants are found in studies of children in English and other language environments. In terms of place of articulation, Irwin's (1947*a*) longitudinal study of English-learning children reported that the frequency of velar and glottal consonants dramatically decreases by the end of the first year, when dental and labial stops [d, b, m] appear to be the second-most frequent consonants after [h]. Kent & Bauer (1985) also found that bilabial and apical consonants occurred most often in the consonants produced by infants aged 1;1 in various utterance patterns (e.g. CV, VCV and CVCV).

Regarding manner of articulation, Irwin (1947*b*) indicated that after age 0;6 the frequencies of [h] and [ʔ] decrease gradually, when various stops and nasals emerge around the time that the velum and epiglottis separate (Kent & Vorperian, 1995). Similarly, stops, nasals and fricatives dominate in consonants produced by infants aged 1;1 (Kent & Bauer, 1985). Among these, stops represent almost three-fourths of the total consonant production in the major syllabic pattern (CV). Although prevocalic consonants (in CV) tend to be frontal stops, postvocalic consonants (or consonantal endings) and intervocalic consonants tend to be continuants (fricatives, nasals in VC, and fricatives, glides and liquids in VCV) (Kent & Bauer, 1985; Stoel-Gammon, 2002). Furthermore, the majority of prevocalic stops in CV syllables are voiced (voiced/voiceless ratio = 3:1; Kent & Bauer, 1985). The distribution of consonants in terms of place, manner and voicing suggests that the voiced bilabial and apical stops [b, d] are the most frequent prevocalic consonants in English-learning infants' vocalizations (Kent & Bauer, 1985).

The developmental trends just reviewed for English-learning children can also be seen in children from other language environments. These apparent universal trends are: (1) labials and dentals occur more often than velars; and (2) stops and nasals occur more often than fricatives, liquids or glides.

In Mandarin-learning children, the developmental order of front and back consonants remains unclear. Similar to data on the production of early words in English-learning infants, labial and dental stops are generally acquired before velar stops (Zhu & Dodd, 2000; Hsu, 2003), but back consonants (e.g. [k, k^h, x] before [f, l, ts^h]) have also been reported to emerge around age 1;6 (Jeng, 1979). This discrepancy in findings was not attributed to consonant distribution patterns in adult language, in which velars occur as often as labials and less often than alveolars (Cheng, 1982). Further research is needed on both babbling and early speech to clarify these contradictory findings. For various manners of articulation, a universal developmental order was also found in Mandarin data. That is, nasals are acquired (age 0;11 to 1;8) before fricatives (1;8 to 2;0) and affricates (after 2;2) (Yue-Hashimoto, 1980; Zhu & Dodd, 2000).

Language-specific characteristics of consonants. In addition to universal trends in consonant development, consonant production by children in different linguistic environments varies significantly by individual place and manner category. Language-specific characteristics of consonant production can be found as early as age 0;9 to 0;10. That is, language-specific consonant production is found at an age similar to that for language-specific vowel production (de Boysson-Bardies & Vihman, 1991; de Boysson-Bardies, Vihman, Roug-Hellichius, Durand, Landberg & Arao, 1992).

The consonant production of infants from four different linguistic environments (French, English, Swedish and Japanese) consistently showed language-specific distribution patterns throughout the 0- to 25-word stages (roughly age 0;9 to 1;5) in both babbling and early words (de Boysson-Bardies *et al*., 1992). French-learning infants produced significantly more labials than Swedish- and Japanese-learning infants, and French-learning infants produced velars the least frequently of all four groups. Furthermore, Swedish-learning infants produced stops significantly more often and nasals significantly less often than French-learning infants.

Language-specific characteristics are also found in Mandarin-learning children. Before reviewing consonant development in Mandarin-learning infants, the Mandarin consonant system is reviewed to provide essential background information. Mandarin has 21 consonants (initials) plus [ŋ], which occurs only in coda position (Luo, 1992; Table 2). Aspiration, a distinctive feature in Mandarin, is used to differentiate six minimal pairs: [p/p^h, t/t^h, k/k^h, tʂ/$tʂ^h$, ts/ts^h, tɕ/$tɕ^h$]. The three series of affricates and fricatives, [tʂ, $tʂ^h$, ʂ], [ts, ts^h, s] and [tɕ, $tɕ^h$, ɕ], are analyzed in different ways.

https://doi.org/10.1017/S0305000909009581 Published online by Cambridge University Press

TABLE 2. *Consonants in Mandarin*

Manner / Place	Stop		Nasal	Fricative		Affricate		Lateral
	Unasp	Asp		Vls	Vd	Unasp	Asp	
Labial	p	pʰ	m					
Labio-dental				f				
Alveolar	t	tʰ	n	s		ts	tsʰ	l
Retroflex				ʂ	ʐ	tʂ	tʂʰ	
Alveo-palatal				ɕ		tɕ	tɕʰ	
Velar	k	kʰ	(ŋ)	x				

Some researchers regard them as allophonic variations with complementary distributions, but others treat them as an independent series. The series [tʂ, tʂʰ, ʂ, ʐ] has lost its retroflex quality and is realized as [ts, tsʰ, s, z] in daily Mandarin conversations spoken in Taiwan. Another sound change in Mandarin spoken in Taiwan is that postvocalic [ŋ] is frequently replaced by [n] (Tse, 1992).

Reanalysis of consonants in a database with over 900 000 syllables (Liu *et al.*, 1975; Cheng, 1982) shows that alveolars (including alveolars, retroflexes, alveo-palatals) occur significantly more often (73·3%) than either labials (13·62%) or velars (13·08%). In the current study, comparison of infants' vocalization and child-directed speech was facilitated by grouping consonants into three major categories: labials, alveolars and velars. ALVEOLARS are used as a general category that includes dentals, post-alveolar retroflexes and palatals. Since infants lack mature dentition, the three pairs of unaspirated/aspirated affricates and fricatives in the adult system are grouped as one category (alveolar) because they share one property – raising the tongue above its presumed rest or neutral position. Moreover, ALVEOLAR is used for ease of comparison with previous studies. As in English and most other languages, stops occur most frequently in Mandarin (34·32%) among the five manners of articulation. One major characteristic of Mandarin is the high frequency of affricates (26·89%), which is slightly higher than that of fricatives (24·97%). Nasals (7·6%) and laterals (6·22%) are the least frequently used consonants in Mandarin.

In contrast to the universal acquisition order of front to back consonants, velar fricatives [x] are acquired earlier than dental [s, ʂ] and palatal fricatives [ɕ] (Jeng, 1979; Zhu & Dodd, 2000), although velar fricatives seem to occur much less often than dental and palatal fricatives in adult language (Cheng, 1982). Among all three nasals in Mandarin, the labial nasal [m] is acquired first. Of the other two nasals, alveolar nasal [n] seems to be less stable than velar nasal [ŋ] in the developmental process. It was reported that [ŋ] is stabilized (1;6–2;0) before [n] (2;1–2;6), and syllable-final [n] is replaced by [ŋ] (Zhu & Dodd, 2000). The late stabilization of [n] in the consonant production of Mandarin-learning children is different from that in English-learning infants and merits further study. Late acquisition of [n] is also found in Cantonese-learning children (Wong & Stokes, 2001).

The acquisition of affricates in Mandarin argues against Jakobson's implicational rule. Mandarin-learning children acquire affricates [ts, tɕ] earlier than [s, ɕ] (Hsu, 2003). Before affricates are completely acquired, they can be replaced either by dental stops [t] or by other affricates (e.g. [ts] substitutes for [ts^h, $tɕ^h$]), but they have never been found to be replaced by fricatives (Zhu & Dodd, 2000; Zhu, 2002). On the other hand, fricatives are sometimes replaced by affricates. For example, the affricate [tɕ] has been substituted for fricatives [s, ʂ, ɕ] (Zhu & Dodd, 2000; Zhu, 2002; Hsu, 2003). This error pattern is also found in Cantonese-speaking children (So & Dodd, 1995). The production of more affricates than fricatives by Mandarin-learning children may relate to the high frequency of affricates in Mandarin.

Developmental continuity

Both developmental discontinuity and continuity have been reported for children in different language backgrounds.

Developmental continuity of vowels. The relationship between the vowel production systems in babbling and early speech was systematically examined in a longitudinal study of a single child from age 1;2 to 1;8 (Davis & MacNeilage, 1990). However, this comparison of vocalic production in babbling and early speech may have been biased by the sampling of vocalization types in Davis and MacNeilage's data, which reflected four times as much word production as babbling between 1;2 and 1;8, perhaps failing to capture the transition in vowel production. Studies on younger age groups, covering the transition from mostly babbling to increasing production of lexical items, may provide a more complete picture of the continuity of vowel development. The current study addressed this issue by examining vocalic production between 0;7 and 1;6, from babbling to early speech.

Developmental continuity of consonants. The phonetic systems of babbling and early speech in English-learning infants resemble each other in various aspects of consonant development. First, the frequently occurring consonants – stops and nasals [p, b, d, k, g, m, n, ŋ] – in early words (produced from 0;10 to 1;2; Waterson, 1978) are a direct outgrowth of the most frequent sounds [b, t, d, g, m, ʔ] in babbling (Irwin, 1947*a*, 1947*b*), except for [h], which is produced frequently only in babbling (Mowrer, 1980). Second, besides this general similarity, the consonant production patterns for each child are consistent in both babbling and early words (Vihman, Macken, Miller, Simmons & Miller, 1985).

Third, initial stops and final fricatives occur with high frequency in both early words and babbling. This unequal distribution of initial stops and fricative endings in early words is shown by the substitution of stops for initial fricatives (e.g. [top] for [sop]) and of fricatives or affricates for final stops (as in [bæs] for [bæk]; Mowrer, 1980). This unequal distribution also reflects sound patterns in babbling, in which stops occur more often than fricatives in initial position, and conversely fricatives (or continuants) occur more often than stops in final position (Kent & Bauer, 1985). Fourth, developmental continuity has also been found in aspiration, voicing and gliding. For example, the deaspiration of initial stops in early words (e.g. [bɛt] for [pɛt]) is an outgrowth of the high frequency of unaspirated initial stops in babbling (Mowrer, 1980).

Fifth, in contrast to the prominence of voiced stops in initial position, consonants in final position tend to be voiceless stops, fricatives or nasals in both early words and babbling, as seen in the devoicing of final consonants in early words (e.g. [bʌk] for [bʌg]; [fʌs] for [fʌz]) and the high frequency of voiceless final consonants in babbling (Mowrer, 1980; Kent & Bauer, 1985). Sixth, substitution of glides [w, j] for prevocalic liquids [l, r] in early words correlates with a higher frequency of prevocalic glides than liquids in babbling (Mowrer, 1980). Seventh, children tend to reduce consonant clusters into single consonants (e.g. [ti] for [tri]) and to delete final consonants in keeping CV syllable structures in early words (Mowrer, 1980). This developmental trend likely results from the higher frequency of CV syllables than of VC syllables and the scarcity of consonant clusters in babbling (Kent & Bauer, 1985). The parallel patterns of consonant production in babbling and early words reviewed above strongly support developmental continuity. However, phonetic development comprises both continuous and discontinuous processes.

As to the relative frequency of labials and alveolars, 'labial regression' is considered counter-evidence of continuity in consonant production between babbling and early words (MacNeilage, Davis & Matyear, 1997). Three of the four children in MacNeilage *et al.*'s longitudinal study produced more alveolars in babbling, whereas three of them produced more labials in early words. That is, the ratio of labials to alveolars in early words (range = 2·04–2·86) was greater than that in babbling. This high alveolar frequency in babbling was interpreted as reflecting bias from the ambient language. Also, the regression to labials, which are easier to produce in early words, was thought to facilitate the variegation of vowel production in early words. However, Vihman *et al.* (1994) found that labials are produced more often than alveolars in child-directed English speech. Thus, 'labial regression' may reflect the phonetic patterns of infants' ambient language and the argument of 'labial regression' should be examined with data from other language groups and more subjects.

In sum, since most studies on vowel development have focused either on babbling in the first year or word production in the second year and later, developmental continuity can be best approached by empirical studies that cover a broader age range, from late babbling to the second year. Moreover, verification of early evidence for language-specific vowel production (de Boysson-Bardies *et al.*, 1989; de Boysson-Bardies, 1993) requires additional studies of infants learning other languages and younger than 0;10. These two issues are addressed by the current study.

Various aspects of consonant development in Mandarin-learning children differ from the general pattern of consonant development in other languages. These discrepancies can be attributed partly to consonant distribution patterns in the ambient language and dialectal influence, but more studies are needed to clarify these discrepancies and to trace the earliest time of language-specific consonant production. In particular, using larger samples would control for bias from individual differences and including younger age groups would allow comparison of the frequencies for various consonant categories in babbling and in early words.

Further studies on vowel and consonant development from babbling to early word production from less-studied languages might confirm the general patterns found so far and resolve contradictory findings on developmental order in babbling and early speech. Therefore, this cross-sectional study with Mandarin-learning infants investigated language-specific patterns and developmental continuity between babbling and the production of first words by asking two questions:

(1) Do vocalic and consonantal developmental patterns in Mandarin-learning infants resemble those of other language groups (reflecting underlying universal constraints) or do these patterns reflect the pattern of their major linguistic input?
(2) Do the patterns of vowel and consonant distribution observed at the early stages persist through later stages of infant vocalizations?

METHODS

Design and subjects

Twenty-four infants, recruited by informal referral from the community surrounding Tainan City (Taiwan), were divided into two age groups: G1 (12 infants, 0;7 to 1;0, representing mostly babbling vocalizations) and G2 (12 infants, 1;1 to 1;6, mostly words). In addition to infant vocalizations, caregivers' speech in caregiver–child interactions was analyzed to study the phonetic characteristics of the infants' language environments. For subjects' characteristics, see Table 3.

Parents and caregivers used Mandarin when interacting with infants. Although direct evidence for developmental patterns might not be shown

TABLE 3. *Subject characteristics* (N=24)

Subject	Gender	Age (year, month, day)	Number of CV syllables	CV syllables/all syllables	Vocabulary produced (words)
S1	M	0;6.28	37	0·2	1
S2	M	0;7.22	73	0·2	1
S3	F	0;9.07	117	0·5	11
S4	F	0;9.18	66	0·2	5
S5	M	0;9.22	129	0·6	10
S6	M	0;10.06	73	0·4	0
S7	F	0;11.00	119	0·5	3
S8	F	0;11.00	163	0·6	3
S9	M	0;11.18	279	0·9	13
S10	F	1;0.00	137	0·7	9
S11	M	1;0.20	262	0·8	0
S12	M	1;0.25	42	0·2	2
S13	F	1;2.08	141	0·5	60
S14	M	1;3.19	204	0·6	59
S15	M	1;3.24	309	0·7	65
S16	F	1;4.04	224	0·6	47
S17	M	1;4.19	88	0·5	11
S18	F	1;4.23	179	0·6	60
S19	F	1;5.00	438	0·7	76
S20	F	1;5.02	649	0·7	73
S21	M	1;5.11	590	0·8	62
S22	F	1;5.15	354	0·8	87
S23	M	1;5.18	169	0·5	5
S24	M	1;5.23	149	0·3	16

by cross-sectional data, these data provide a reliable profile of development when the design is carefully controlled as described below.

Data collection and analysis

Infants' vocalizations and child-directed speech were audio-recorded with a DAT recorder (Sony TCD-D8) and a wireless lapel microphone (Telex ProStar R-10) while observing their natural daily activities at home or in daycare centers. Infants' spontaneous vocalizations and adults' child-directed speech were transcribed, compiled and analyzed for frequencies of the major categories of consonants and vowels.

Data were selected by six criteria. First, to incorporate all possible precursors of speech in infants' vocalizations, 'speechlike' sounds were broadly defined to exclude only vegetative or reflexive sounds (e.g. cries, coughs, breathing noises). Second, among all speechlike vocalizations, only sequences with at least one C-like and one V-like sound (canonical syllables) were analyzed. In identifying canonical syllables, perceptual judgments about well-formed syllables were based on previous criteria documented in the literature (e.g. Davis & MacNeilage, 1990; Vihman, 1992). That is, isolated vowels and sequences of vowels preceded or followed by glottal stops were not included. Third, all spontaneous productions were transcribed and analyzed, while imitated data were not. Fourth, each CV syllable in isolated syllables, reduplicated babbles or variegated babbles were analyzed as individual CV syllables. Fifth, no distinction was made between babbling and early words, and infants' CV productions were not analyzed for accuracy of attempted adult words. The rationale for this decision is the difficulty in establishing objective criteria for reliably identifying the target meaning of early vocalizations with very limited contextual information from only audio recordings. Sixth, all spontaneous CV sequences produced during all 24 infants' 45-minute recordings were phonetically transcribed. For longer recordings, the 45-minute section with the most CV syllables was transcribed and analyzed.

Median differences in the occurrence (frequency) of categories of vowel and consonant productions were compared by one-sample Wilcoxon tests. Median differences were examined rather than mean differences because the former are less influenced by outliers and are more valid measures of patterns from small samples with large individual differences. However, mean values are included to confirm distribution patterns. The distribution patterns of vowel and consonant production between infants' and adults' systems in child-directed speech (language-specific characteristics in research question one) and between early and later developmental stages (research question two) were compared by two-sample Wilcoxon tests. Correlations between variables were determined by Spearman's r. The rationale for choosing and dividing the age range 0;7–1;6, the criteria for excluding extreme individual differences, the data collection procedure, the data analysis methods and the statistical analysis have been described (Chen & Kent, 2005).

Early CV utterances, which emerge between 0;6 and 1;0, were emphasized in this study because of their critical role in early phonetic development. CV syllables occur second-most frequently after single vowels and diphthongs (Kent & Bauer, 1985). Among all syllable types, CV syllables are believed to represent children's phonetic capabilities. This hypothesis is based on two observations (Kent & Bauer, 1985): (1) CV syllables comprise 47·6% of all syllables; and (2) consonant and vowel distribution patterns in other frequently occurring utterance patterns (e.g. V, CVCV) parallel the patterns in CV syllables.

Intra-transcriber reliability was checked by re-transcribing a randomly selected 5-minute sample (11% of the total) from each of the 24 infants' recordings. Inter-transcriber reliability was checked by having an experienced, monolingual English-speaking transcriber re-transcribe a 5-minute sample from each infant's data. This transcriber received only instruction for broad transcription without introducing the Mandarin phonological system and without knowledge of grouping all low vowels as low-central vowels. Estimates of intra-transcriber and inter-transcriber reliabilities for major vowel categories were over 95% and 70% (major confusion found between [ɛ, ə]), respectively. Intra-transcriber reliability with 1,245 CV syllables for vowel backness was 95·4% and for vowel height was 95·5%. Inter-transcriber reliability with 1,052 CV syllables for vowel backness was 78·1% and for vowel height was 71·6%. It is worth noting the high degree of agreement between transcribers in identifying the low vowel [a] (intra- and inter-transcriber reliabilities of 98·2% and 86·2%, respectively) because the prominence of [a] in Mandarin-learning infants was a major finding in this study. Intra-transcriber reliabilities were 99·8% and 99% for consonant place and manner of articulation, respectively. Inter-transcriber reliabilities were 98·8% and 97·5% for consonant place and manner of articulation, respectively.

RESULTS

Vowels

Vowels were analyzed in terms of height (high, mid, low) and backness (front, central, back). Since this study's goal was to investigate the development of major categories of vowels in CV syllables, broad transcription was used, based on the IPA system. Furthermore, single vowels occurring with high frequency and with high intra- and inter-transcriber reliability were also noted. Although the V-like sounds in early infant syllables do not exhibit the same qualities as in adults' speech, the term VOWELS is used in this study. This categorization enabled comparison with findings of other studies without denoting the level of maturity in the production of segmentals in early syllables.

The most frequently occurring vowels in infants' vocalizations in this study were [ɛ, ə, a]. The only low vowel in Mandarin spoken by adults is [a], which was categorized as a low-central vowel in analyzing both infants' and caregivers' data.

Vowel production in infant vocalization and child-directed speech

Analysis of the frequency distribution of vowel backness shows parallel patterns in infants' vocalizations and caregivers' child-directed speech (Table 4). That is, central vowels occurred significantly more often than front vowels, and front vowels occurred significantly more often than back vowels. These two sets of data did not differ significantly in front vowel production, but did differ in central and back vowels. More central vowels were produced in infants' vocalization than in child-directed speech, due to

TABLE 4. *Comparison of vocalic production by major category in infants and caregivers* (N=48)

Vowel category[a]	Infants			Caregivers				
	Mean	SD	Median	Mean	SD	Median	r^b	$t_w{}^c$
Front	29·84	21·84	30·20	31·59	4·97	33·06	0·05	168·0
Central	62·85	26·02	63·46	47·51	4·63	46·92	0·26	234·0*
Back	7·31	12·48	1·35	20·90	6·79	20·15	0·08	271·0*
High	13·08	14·89	8·70	42·53	6·14	43·05	0·50*	336·0*
Mid	40·50	23·52	38·60	27·04	5·92	27·50	0·22	698·5
Low	46·42	29·43	40·15	30·43	7·65	28·85	0·54*	676·0

* $p < 0·017$ (0·05/3).
[a] The sum of mean percentages within each category is 100.
[b] Spearman's r (directional).
[c] Two-sample Wilcoxon test (non-directional).

the high frequency of low-central vowels [a] in infants' data. Moreover, fewer back vowels and vowel varieties were found in infants' utterances.

Unlike the similar vowel backness patterns seen in infants and adults, their vowel height patterns differed. In child-directed speech, high vowels occurred most frequently among the three major vowel categories. Conversely, in infants' vocalizations, both low vowels and mid vowels occurred significantly more often than high vowels. High vowels occurred significantly more often in adults' child-directed speech than in infants' utterances. However, the percentages of high vowels produced by individual children correlated with the percentages produced by their respective caregivers, as shown by the significant correlation. Low vowels had the highest correlation between infants' and caregivers' data (Table 4).

The sounds in each vowel category (high, mid, low, front, central, back) were collapsed in both cases above and broad transcription was used, whereas in the following analysis they were separate. The analyses further identified these sounds by vowel category in a vowel quadrilateral (Table 5): high front [i, y], high back [u], mid front [ɛ], mid central [ə], mid back [ɔ] and low central [a]. Other vowels in Table 1 (e.g. high-central vowel) were not found in the data at this stage. Except for the late-acquired vowels, this analysis revealed several patterns. First, the low vowel [a] occurred significantly more often than the high-front vowels [i, y] in both infants' vocalizations and child-directed speech. Second, while adults used the high-front vowels [i, y] significantly more often in child-directed speech than the mid-front vowel [ɛ], the mid-front vowel occurred in infants' vocalizations significantly more frequently than the high-front vowel. The mid-front vowel [e] in the adult system is acquired late. Before fully developing [e], [ɛ] is used often in infants' speech. Third, the frequency distribution of

https://doi.org/10.1017/S0305000909009581 Published online by Cambridge University Press

TABLE 5. *Comparison of vocalic production by vowel subcategory in infants and caregivers* (N=48)

Vowel subcategory	Infants (%)			Caregivers (%)				
	Mean[a]	SD	Median	Mean[a]	SD	Median	r^b	$t_w{}^c$
High front	10·83	12·96	7·90	25·49	4·87	26·05	0·20	372·5*
High central	0·15	0·55	0·00	3·18	2·22	2·40	0·42	315·0*
High back	2·06	4·93	0·00	13·85	6·44	13·00	0·34	328·0*
Mid front	19·01	13·37	18·10	6·10	2·59	5·70	−0·06	759·0*
Mid central	16·25	19·78	8·40	13·90	5·49	14·80	0·16	518·5
Mid back	5·28	10·85	0·15	7·05	4·08	5·90	0·16	441·5*
Low central	46·42	29·43	40·15	30·43	7·65	28·85	0·54*	676·0

* $p < 0{\cdot}006$ (0·05/9).
[a] The sum of the mean percentages in each column is 100.
[b] Spearman's *r* (directional).
[c] Two-sample Wilcoxon test (non-directional).

TABLE 6. *Comparison of major vowel categories produced by two infant age groups* (N=24)

Vowel subcategory	G1 infants (ages 0;7 to 1;0) (%)			G2 infants (ages 1;1 to 1;6) (%)			
	Mean[a]	SD	Median	Mean[a]	SD	Median	$t_w{}^b$
Front	20·29	20·72	15·79	39·40	19·21	36·24	114*
Central	78·91	21·18	81·46	46·79	20·11	45·51	97*
Back	0·80	1·23	0·00	13·82	15·22	10·77	103*
High	4·34	6·47	0·00	21·81	15·97	20·95	99·5*
Mid	38·78	30·35	32·20	42·23	15·13	41·75	139·0
Low	56·88	34·44	62·20	35·96	19·65	33·00	178·0

* $p < 0{\cdot}017$ (0·05/3).
[a] The sum of mean percentages within each category is 100.
[b] Two-sample Wilcoxon test comparing medians (non-directional).

infant vs. adult vowel production by category in the vowel quadrilateral were significantly different, except for the mid-central vowel [ə] and the low vowel [a] (Table 5). Fourth, only the low vowel [a] was significantly correlated between the infants' and caregivers' data (Table 5). In conclusion, infants' vowel production more accurately reflected caregivers' child-directed speech in terms of vowel backness rather than vowel height, and for the frequently occurring vowel [a].

Developmental changes in vowel production

In general, the three vowel categories (front, central and back) display similar patterns between G1 and G2 (Table 6). At least half the vowels produced by infants were central vowels (Table 6), mostly due to the high frequency of the low-central vowel [a]. Of the 12 G1 infants, 9 already showed this language-specific characteristic of the low-central vowels, some as young as age 0;7.

Infants' vowel categories changed between ages 0;7 and 1;6 in three major ways. First, the median frequency of front vowels increased from 16% in G1 to 36% in G2. Second, the median frequency of central vowels significantly decreased from 81% in G1 to 46% in G2. Third, the median frequency of back vowels significantly increased from 0% in G1 to 11% in G2. These developmental changes led to the infants' vowel productions gradually resembling the adult model in child-directed speech in vowel backness, as seen in the frequency distribution of vowels in G2 (Table 6) and in the caregivers' data (Table 4). The frequency distribution pattern in G2 infants resembles those of caregivers (G2: front 36%, central 46%, back 11%; caregivers: front 33%, central 47%, back 20%).

Regarding the vowel height of infant vocalizations, low and mid vowels occurred more frequently than high vowels in both G1 and G2 (Table 6). Similarly, these two age groups did not differ significantly in the frequency of mid and low vowels, but the occurrence of high vowels increased significantly (Table 6). Over the course of development, the median frequency of high vowels increased from 0% in G1 to 21% in G2. This developmental trend was also found in the data of individual infants. In G1, 8 of 12 infants produced low vowels more frequently than either high or mid vowels, whereas in G2, low vowels occurred most frequently for only three infants. Additionally, high vowels were never produced more frequently than mid or low vowels by any infant in G1, whereas high vowels were produced more frequently than mid or low vowels by three infants age 1;5 (G2). Additionally, the variability in producing the early developed mid and low vowels declined greatly from G1 to G2 (shown by the decreases in standard deviations, Table 6). These findings and the general patterns of G1 show a developmental trend in G2 of progressing toward the adult model (G1: high 0%, mid 32%, low 62%; G2: high 21%, mid 42%, low 33%; caregivers: high 43%, mid 28%, low 29%).

Analysis of vowel production via the nine vowel categories in the vowel quadrilateral revealed several similarities between G1 and G2 (Table 7). First, the low vowel [a] occurred most often in both G1 and G2. All infants produced [a] with high frequency. This language-specific characteristic could be found even among infants age 0;7. In fact, the low vowel [a] occurred most often in the majority of infants studied. Second, the prominent pattern of the mid-front [ɛ] and mid-central [ə] vowels occurring frequently in G1 persisted in G2. Third, the low vowels and mid-front vowels occurred more often than the high-front vowel. Similarly, the

TABLE 7. *Comparison of vocalic production by vowel subcategory in two infant age groups* (N = 24)

Vowel subcategory	G1 infants (ages 0;7 to 1;0) (%)			G2 infants (ages 1;1 to 1;6) (%)			
	Mean[a]	SD	Median	Mean[a]	SD	Median	t_w[b]
High front	4·05	6·22	0·00	17·61	14·56	15·55	101·5*
High central	0·00	0·00	0·00	0·30	0·76	0·00	—
High back	0·30	0·80	0·00	3·83	6·59	1·20	106·5
Mid front	16·25	17·25	11·95	21·78	7·70	18·40	127·5
Mid central	22·03	25·32	14·35	10·47	10·22	6·45	168·0
Mid back	0·79	0·97	0·00	10·05	14·00	6·35	106·5
Low central	56·88	34·44	62·20	35·96	19·65	33·00	178·0

* $p < 0{\cdot}006$ (0·05/9).
[a] The sum of mean percentages in each column is 100.
[b] Two-sample Wilcoxon test comparing medians (non-directional).

frequency of corresponding vowel categories did not differ significantly in G1 and G2 (except for the high-front vowel, as described below).

Apart from these aspects of developmental continuity between G1 and G2, some developmental changes occurred. From G1 to G2, the median frequency of the high-front vowel increased from 0% to 16% ($p < 0{\cdot}006$). Although the frequencies of other vowel categories did not differ significantly, the median frequency of the low vowel decreased from 62% to 33%. This decrease might have been due to a greater variety of vowel categories in G2 than in G1. In fact, all vowel categories in child-directed speech were observed in G2.

In summary, the general patterns of vowel production seen in G1 continued in G2. These included high frequencies of low and mid vowels (especially the mid-front and mid-central vowels). Developmental changes were also seen in G2 in the increased frequency of front and high vowels (especially the high-front vowel). Other developmental changes included an increased frequency of high and mid-back vowels. Moreover, variability in the frequency of early developed vowel categories declined in G2.

Consonants

Consonants were analyzed for frequency of consonantal categories by place and manner of articulation. Although consonant-like sounds in early syllables do not exhibit exactly the same qualities as consonants in adults' speech, they were categorized as adult-like consonants in this study. To investigate the development of the major consonant categories in CV syllables, broad transcription was used, based on the IPA system.

https://doi.org/10.1017/S0305000909009581 Published online by Cambridge University Press

TABLE 8. *Comparison of consonant production by major category in infants and caregivers* (N=48)

Major consonant category	Infants			Caregivers				
	Mean[a]	SD	Median	Mean[a]	SD	Median	r^b	$t_w{}^c$
Labial	43·45	33·49	36·85	21·69	7·20	20·35	0·37	688·5
Alveolar	44·75	31·30	44·75	57·84	5·94	58·85	0·22	507·0
Velar	11·80	18·02	2·60	20·47	5·51	19·50	0·31	426·5*
Stop	69·15	27·75	70·20	34·34	6·73	32·65	0·32	807·5**
Nasal	23·11	26·61	11·00	15·21	5·10	15·75	0·56*	579·0
Fricative	4·67	11·73	0·00	18·91	4·09	18·40	0·34	335·0**
Affricate	3·06	5·59	0·30	19·78	4·27	21·25	0·13	319·0**
Lateral	0·01	0·06	0·00	11·76	4·16	11·45	−0·14	300·0**

* $p<0·017$ (0·05/3). ** $p<0·01$ (0·05/5).
[a] The sum of the mean percentages within each category is 100.
[b] Spearman's r (directional).
[c] Two-sample Wilcoxon test (non-directional).

Consonant production in infants' vocalizations and child-directed speech

Place of articulation. Analysis of the pooled data for G1 (0;7 to 1;0) and G2 (1;1 to 1;6) reveals that alveolars and labials occurred more frequently than velars, although this difference was not significant. In addition, most G1 and G2 infants consistently produced far fewer velars than alveolars and labials (Table 8).

In caregivers' data, alveolars occurred significantly more often than labials or velars (Table 8). Comparison of infants' and caregivers' data reveals several patterns. First, alveolars in both sets of data have similar medians (range=45–59%). Second, the ratio of median labial frequency to median alveolar frequency in the infants' data (0·82) was much larger than in the adults' data (0·35). That is, in relation to alveolars, labials were produced much more frequently by infants than by adults. However, labials were not produced more often than alveolars by infants or caregivers. Third, the frequency of labials and alveolars produced by infants and caregivers was not significantly different. Fourth, velars occurred significantly more often in caregivers' than in infants' speech. In general, the place of consonant articulation in infants' vocalizations showed a frequency distribution similar to that of child-directed speech, especially for alveolars and labials.

Manner of articulation. Stops occurred in infants' vocalizations significantly more often than the other four consonant categories (Table 8). Nasals occurred second-most frequently and significantly more often than the other three categories. Furthermore, all infants produced relatively fewer fricatives, affricates and laterals than stops and nasals. Although infants produced alveolars significantly more often than velars, they produced the

TABLE 9. *Comparison of consonant production by major category in two infant age groups* (N = 24)

Major consonant category	G1 infants (aged 0;7 to 1;0) (%)			G2 infants (aged 1;1 to 1;6) (%)			
	Mean[a]	SD	Median	Mean[a]	SD	Median	t_w[b]
Labial	58·93	38·44	54·10	27·97	18·56	26·95	183·0
Alveolar	39·15	38·02	37·95	50·34	23·11	49·70	138·0
Velar	1·92	3·05	0·40	21·69	21·40	15·90	97·0*
Stop	70·00	33·23	87·30	68·30	22·45	61·55	158·0
Nasal	28·56	33·13	11·00	17·66	17·84	14·25	164·0
Fricative	0·30	0·85	0·00	9·00	15·66	2·25	103·0**
Affricate	1·14	2·82	0·00	5·01	7·01	2·30	109·0
Lateral	0·00	0·00	0·00	0·03	0·09	0·00	—

* $p < 0·017$ (0·05/3). ** $p < 0·01$ (0·05/5).
[a] The sum of the mean percentages within each category is 100.
[b] Two-sample Wilcoxon test (non-directional).

velar fricative [x] more often than alveolar fricatives [s, z, ɕ] but the difference was not significant.

In the adults' child-directed speech, stops were used most frequently and occurred significantly more often than the other four consonant categories. However, affricates and fricatives were produced by adults significantly more often than nasals. This pattern contrasted with that in the infants' data. The frequency distributions for infants' and adults' manner of consonant articulation differed significantly except for nasals, whose frequency distributions were significantly correlated (Table 8).

These two sets of data on consonant place and manner suggest that infants' vocalizations reflect adults' child-directed speech in the categories of alveolars, labials, stops and nasals. However, infants' velars, fricatives, affricates and laterals were still in the early stages of development, as shown by their relatively infrequent appearance.

Developmental changes in consonantal production

Place of articulation. Labials and alveolars were produced significantly more often than velars by infants in groups G1 and G2 (Table 9). Nevertheless, these two groups differed in several aspects. First, velars were produced significantly more often in G2 than in G1. Second, the frequencies of labials and alveolars were not significantly different in G1, whereas in G2 alveolars occurred significantly more often than labials. Third, the ratio of the labial median to the alveolar median was much larger in G1 (1·43) than in G2 (0·54), indicating that, in relation to alveolars, labials were produced much more frequently by G1 infants than by G2 infants (Table 9).

https://doi.org/10.1017/S0305000909009581 Published online by Cambridge University Press

In general, G1 and G2 showed similar frequency distribution patterns with respect to place of consonant articulation. Despite this general continuity between G1 and G2, however, developmental changes across these age periods progressed toward the adult model. Unlike the G1 articulation data, those of G2 infants resemble those of caregivers in both frequency distribution (G2: labials 27%, alveolars 50%, velars 16%; caregivers: labials, 20%, alveolars 59%, velars 20%) and labial to alveolar ratios (G1: 1·43, G2: 0·54, caregivers: 0·35) (Tables 8 and 9).

Manner of articulation. The frequency patterns for manner of consonant articulation in G1 and G2 showed similar distributions. First, over 60% of the consonants were stops in both groups (Table 9). In fact, stops occurred significantly more often than the other four consonant categories in both G1 and G2. Second, nasals occurred second-most frequently in G1 and G2. Third, infants in G1 and G2 (but to a lesser degree) produced relatively fewer fricatives, affricates and laterals than stops and nasals. These differences were all significant. The similarities between G1 and G2 reflected no significant differences in stops and nasals (Table 9).

Apart from these similarities, the frequency of fricatives increased significantly from G1 to G2. The frequency of affricates also changed between G1 and G2. Among the G1 infants (0;7 to 1;0), only three produced affricates, all at frequencies under 10%, whereas 10 of the 12 infants in G2 produced affricates. Two infants (S20 and S22) produced affricates at 19% frequency, close to the median frequency in child-directed speech (18·4%). These data indicate an early development of fricatives and affricates, with frequencies gradually increasing from G1 to G2.

In general, consonant production in G2 continued the pattern found in G1. That is, labials and alveolars occurred more often than velars, and stops and nasals were produced far more often than fricatives, affricates and velars. Besides these aspects of developmental continuity across two age groups, G2 infants were closer than G1 infants to the pattern of caregivers' speech with respect to labial to alveolar ratios and greater production of fricatives and affricates. Furthermore, G2 infants more consistently produced the major and early developed consonant categories (e.g. labials, alveolars, stops and nasals) than G1 infants, as shown by reduced variability (smaller standard deviations) from G1 to G2 (Table 9).

DISCUSSION

Universal and language-specific patterns

The universal and language-specific characteristics of consonant and vowel production of Mandarin-learning infants in this study are summarized in Table 10.

TABLE 10. *Universal and language-specific characteristics*

	Universal	Language-specific
Vowels	Mid-front and mid-central vowels predominate Low and mid vowels predominate over high vowels	Low vowel [a] predominates over other vowels, comprising 40% of all vowels
Consonants	Labials and alveolars predominate over velars Velar fricatives predominate over alveolar fricatives Stops predominate over other manners Nasals are the second-most frequent manner category	Alveolars predominate over labials Early appearance of prevocalic affricates

Vowels. This study confirms at least two universal characteristics in the development of vowels. First, infants predominantly produced the mid-front vowel [ɛ] and mid-central vowel [ə], resembling the pattern of early vowel production reported for English-learning infants in studies using phonetic transcription. For example, mid-front and mid-central vowels constitute over 70% of vowels produced during the first year (Irwin, 1948) and central and front vowels are produced most frequently by infants aged 1;1 (Kent & Bauer, 1985). The same pattern was found in studies using acoustic analysis. For example, the vowel spaces of non-high front and central vowels become stabilized earlier than any high or back vowels (Buhr, 1980), and vowels produced by infants between 0;3 and 0;9 are mostly mid-front or central vowels (Kent & Murray, 1982).

The second universal characteristic found in this study was the predominance of low and mid vowels over high vowels. In contrast, high vowels were the last-developed vowel category, and their frequency increased only after age 1;0. Accordingly, high vowels had a much lower frequency than mid and low vowels. This finding parallels that of most studies on English-learning infants, using perceptual analysis (e.g. Kent & Bauer, 1985) or acoustic measurement (e.g. Buhr, 1980). The predominance of low and mid vowels (especially [ɛ, ə]) over high vowels is closely related to the anatomical structures of the infant vocal tract. The high frequency of low and mid vowels in infants' vocalizations near the end of their first year reflects an immature system in which the tongue 'moves primarily in an anterior-to-posterior dimension and with little elevation from low carriage' (Kent, 1992: 72).

Apart from these universal patterns, we found that over 40% of vowels produced by Mandarin-learning infants were low vowels [a]. The low vowel [a] was produced most often among all individual vowels. In contrast, both perceptual analyses (Irwin, 1948; Kent & Bauer, 1985) and acoustic analyses (Buhr, 1980; Kent & Murray, 1982) indicate that front [ɪ, ɛ] and central vowels [ə] occur more often than other vowels in English-learning infants. However, the early appearance and acquisition of low vowels found in this study has been reported for Cantonese- or Mandarin-learning infants. For example, the vowel production of Cantonese-learning infants has a lower F2/F1 ratio (i.e. high frequency of the low-back vowel) than that of infants learning English, French or Arabic (de Boysson-Bardies *et al.*, 1989).

In addition, [a] is acquired by age 2;0 in Mandarin-learning children (e.g. Li, 1978; Yue-Hashimoto, 1980; Hsu, 2003). The predominance of [a] over other vowels in Mandarin-learning infants is closely related to the pattern in child-directed speech. The occurrence of low-central vowels (i.e. [a]) in over 28% (highest frequency among all vowels) of adult child-directed speech was an obvious and consistent pattern in this study and in Cheng's (1982) written materials. This language-specific prominence of [a] was found in our subjects before age 1;0 (G1). In fact, the high frequency of [a] occurred even in our youngest subjects (aged 0;7). Thus, our results suggest that language-specific vowel production appears before age 0;10 as reported (de Boysson-Bardies *et al.*, 1989; de Boysson-Bardies, 1993).

Consonants. Several universal characteristics were seen in the development of consonants. First, Mandarin-learning infants in this study produced labials and alveolars more often than velars, similar to the pattern for English-learning infants (e.g. Irwin, 1947*a*; Kent & Bauer, 1985). Second, in direct contrast to this general tendency of labials and alveolars to occur more often than velars, we found that the velar fricative occurred more often than the alveolar fricative. The prominence of the velar fricative is neither unique to our subjects nor specific to the Mandarin environment, as it has been found in English-learning infants, especially in early babbling (Irwin, 1947*b*). The early emergence of the velar fricative in Mandarin-learning infants has also been reported; for example, [x] is acquired before [s, ɕ, f] (Jeng, 1979; Zhu & Dodd, 2000).

Third, we found that stops occurred most often among all manner categories in Mandarin-learning infants. This finding is consistent with the predominance of stops found in English-learning infants (Kent & Bauer, 1985; Vihman *et al.*, 1985) and with findings in cross-linguistic studies (e.g. de Boysson-Bardies & Vihman, 1991).

Fourth, we found that nasals are often produced by Mandarin-learning infants (aged 0;7 to 1;6) and at frequencies that even reached adult levels. Similar patterns have been found in other language groups, e.g. increased production of nasals after age 0;5 to 0;6 (Irwin, 1947*b*) and a prominence of nasals next to stops (Kent & Bauer, 1985) for English-learning infants. Our observed pattern is also consistent with the finding that nasals occur second-most frequently (after stops) in the data of infants from French, English, Japanese and Swedish environments (de Boysson-Bardies & Vihman, 1991).

Aside from these universal tendencies, two language-specific patterns were observed in the consonant production of Mandarin-learning infants: the prominence of alveolars over labials and the early appearance of affricates. First, the relative frequency of labials and alveolars in this study differs from the pattern in other language groups. For example, both French- and English-learning infants produce more labials than alveolars, consistent with the pattern of their input languages (de Boysson-Bardies *et al.*, 1992). Similarly, labials are produced more often than alveolars in English-speaking adults' infant-directed speech (Vihman *et al.*, 1994). This pattern of labial to alveolar ratio was also seen in English-learning infants (MacNeilage *et al.*, 1997). That is, although infants produced more alveolars in babbling (labial to alveolar ratio ranged from 0·23 to 1·01), they produced more labials in early words (labial to alveolar ratio ranged from 2·04 to 2·86).

Conversely, alveolars in Mandarin occur much more often than labials in both child-directed speech and written materials (Cheng, 1982). Infants' vocalizations in our study gradually changed to resemble the adult model. The pattern of more labial than alveolar consonants in early babbling (ages 0;7 to 1;0) changed to more alveolars than labials in later babbling (ages 1;1 to 1;6) (G1: labial to alveolar ratio = 1·43; G2: labial to alveolar ratio = 0·54). Through this development, Mandarin-learning infants' data display language-specific characteristics and do not differ from caregivers' data in the frequencies of labials and alveolars and in the prominence of alveolars.

The second language-specific characteristic found in this study is the early appearance of the prevocalic affricates [ts, ts^h, tɕ, $tɕ^h$]. Although fricatives and affricates show similar frequencies in infants' data, the early appearance of affricates relative to other language groups deserves attention. Although our pooled sample of infants produced, on average, less than 5% affricates in their total consonantal vocalizations, most infants aged 1;1 to 1;6 produced some affricates. In fact, some infants produced 19% affricates, similar to the proportion in caregivers' speech. Moreover, these Mandarin-learning infants produced the four alveolar affricates [ts, ts^h, tɕ, $tɕ^h$] more frequently than the alveolar fricatives [s, z, ɕ]. These findings differ from those for English-learning infants. For example, prevocalic affricates [tʃ, dʒ] are not acquired by four-year-olds, based on the spontaneous data of 100 children (Olmsted, 1971, cited in Ingram, 1989).

The early appearance of affricates has also been reported for Mandarin-learning children. For example, affricates [ts, tɕ] appear earlier (ages 1;5 to 1;8) than fricatives [s, ɕ] and are never replaced by fricatives, although fricatives are sometimes replaced by affricates (Hsu, 2003). Similarly, 90% of 129 Putonghua-speaking children produced the affricates [tɕ, tɕʰ] before age 2;0, although they were not stabilized until age 4;0 (Zhu & Dodd, 2000). A similar pattern of affrication was found before age 2;11 in a study of 268 Cantonese-speaking children (So & Dodd, 1995). Finally, infants exposed to K'iche' as the ambient language produced the affricate /ts/ and lateral /l/ with high frequency, similar to the adult language (Pye, Ingram & List, 1987). These results all point to a strong influence of ambient language in determining aspects of early infant vocalizations. However, this influence is difficult to judge from the extant published data, which are sparse for most languages.

The early appearance and acquisition of affricates in Mandarin-learning infants appears closely related to the frequency distribution of affricates in the linguistic input. Affricates clearly occurred more often than fricatives and second-most often after stops in our sample of adult child-directed speech and in written materials (Cheng, 1982). These results highlight the robustness of early perceptual systems that enable infants to detect sound patterns in the ambient language. They also indicate that lingual articulation in infancy is more proficient than might be induced by studies focusing on the most frequently occurring sound patterns in babbling and early vocalizations. This point is considered in detail below.

Developmental continuity

Our study presents the first systematic analysis of developmental continuity in the vocalizations of Mandarin-learning infants during the transition from babbling to early words. We found significant similarities in vowels and consonants produced by infants aged 0;7 to 1;0 and those aged 1;1 to 1;6, as reported (Vihman *et al.*, 1985; Vihman, 1992; Davis & MacNeilage, 1990). In addition to these similarities, quantitative changes were observed in various aspects of phonetic development across the two age ranges. All the universal characteristics (discussed above) found in G1 continued to be prominent in G2. Furthermore, several language-specific characteristics found in the early stage persisted in the later stage. The developmental continuity and changes found in this study are summarized in Table 11.

First, vowel categories were expanded in G2. For example, although central vowels occurred most often among vowels in both G1 and G2, the high frequency of central vowels (especially [a]) in G1 decreased in G2, accompanied by a higher frequency of front vowels [i, ɛ] in G2. Second, individual differences among infants diminished from G1 to G2, shown by the lower G2 variability in median production of the early-developed categories (central, low and mid vowels, especially [a, ɛ, ə]), suggesting their stabilization. This indication of progress in development is consistent with

TABLE 11. *Developmental continuity and changes*

	Developmental continuity	Developmental changes
Vowels	High frequency of low and mid vowels	Vowel categories expand More front and high vowels Stabilization of early developed vowel categories; less variability among infants
Consonants	Labials and alveolars predominate over velars	Some universal characteristics (e.g. labials and stops) decrease from G1 to G2
	Stops predominate over other manners	Language-specific characteristics gradually became obvious in G2 (e.g. affricates produced by most infants, alveolars predominate over labials)
	Nasals are the second-most frequent manner category	Increased variety of consonants, similar to adult model Stabilization of consonants (labials, alveolars, stops, nasals); less variability among infants

other indices of linguistic development: the development of productive vocabulary (mean of 4·8 words in G1 and 52 words in G2), the production of CV syllables (mean of 125 in G1 and 291 in G2) and the canonical babbling ratio (CV syllables/all syllables; mean of 0·5 in G1 and mean of 0·6 in G2).

Consonant development changed across G1 and G2 in four aspects.

(1) The high frequency of some universal characteristics found in G1 decreased in G2 (e.g. labials and stops). This change correlated with increased variety and complexity in development.

(2) Some language-specific characteristics gradually became obvious in the developmental process. For instance, only a few G1 infants produced affricates, whereas most G2 infants produced affricates. Moreover, the increased frequency of alveolars and decreased frequency of labials from G1 to G2 highlights the language-specific pattern of alveolars occurring more often than labials in G2.

(3) Developmental increases in the variety of consonants produced made the infants' vocalizations more congruent with the adult model. Other than those prominent categories (i.e. alveolars, labials, stops, nasals), the frequencies of velars, affricates, fricatives and laterals increased from G1 to G2.

(4) Individual differences among infants decreased from G1 to G2, as shown by less G2 variability in the production of early developed categories (labials, alveolars, stops, nasals), indicating stabilization of these categories. This aspect of progress in development is consistent with other indices of linguistic development: productive vocabulary, number of CV syllables and the canonical babbling ratio.

In conclusion, we found no sharp discontinuity between babbling (roughly corresponding to G1) and the production of first words (roughly corresponding to G2), as proposed by Jakobson (1941/68). Instead, our quantitative data reveal significant similarities in the production of vowels and consonants by G1 and G2 infants. Moreover, 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.

Language-specific production of consonants (the predominance of alveolars over labials and the prominence of affricates) was seen in some infants under age 1;0 and was obvious for infants above age 1;0. Thus, language-specific consonant production in Mandarin-learning infants occurs primarily between 1;1 and 1;6, a few months later than language-specific vowel production, which was found as young as age 0;7. Nevertheless, the timing of language-specific consonant production found in this study needs to be verified by longitudinal and cross-sectional studies of groups with smaller age ranges because our finding (language-specific consonant production emerges between 1;1 and 1;6) contradicts previous findings, i.e. language-specific consonant production emerges between 0;9 and 0;10 (de Boysson-Bardies & Vihman, 1991; de Boysson-Bardies *et al.*, 1992).

The frequency of vowel errors was found in the Memphis Vowel Project to increase when difficulties increased with the consonant context (Pollock, 2002). Thus, vowel development might be correlated with the development of prevocalic consonants in early syllables through CV association patterns. However, the distribution of vowel production in Mandarin-learning infants is generally irrelevant to prevocalic consonants except that front vowels tend to follow alveolars (Chen & Kent, 2005). Further study of potential CV co-occurrence constraints in early lexical production is crucial to verifying our findings.

Implications for models of phonetic development in infancy

Proclivities in infant sound production are no doubt based in part on articulatory constraints, as discussed by Locke (1983). Some sounds occur substantially more often than others in the early phonetic repertoire of babbling, and these sound preferences apply to a variety of languages. Although bilabial consonants appear with high frequency in the early stages of development, evidence of lingual articulation is abundant for both consonants and vowels. The production of alveolar affricates by the infants in this study is one example of early proficiency in lingual articulation. Precocity in tongue muscles would support suckling, therefore having high survival value.

Although molecular and biophysical information is limited on the tongue of the human infant, infants can clearly move the tongue independently of

https://doi.org/10.1017/S0305000909009581 Published online by Cambridge University Press

the mandible (albeit exploiting tongue synergies) and can use the tongue to produce a variety of vowels and consonants. The distribution of these sounds does not closely mirror the ambient language, which is not surprising given the anatomic differences between the infant and adult vocal tracts. To produce even a modest set of consonants, infants must be able to exercise differentiated control over the three-dimensional matrix of the tongue muscle fibers. The data on affricates and low vowels in this study are evidence that the frequency of sounds in the ambient language is a robust influence on the phonetic characteristics of early vocalizations. A theory of phonetic development must be able to accommodate this influence, which is a key feature of language learning.

Babbling, as proposed by the 'Frames, then Content' hypothesis (Davis & MacNeilage, 1990) for speech acquisition, results from the production of syllabic 'Frames' that reflect rhythmic mandibular oscillation. The theory predicts that relatively little of the intrasyllabic and intersyllabic 'Content' of the syllable-like cycles will be under mandible-independent control. Although this hypothesis accounts for some general patterns, such as affiliations of elements in CV syllables, it does not explain many other facets of infant vocalization, including the influence of high-frequency sounds in the ambient language. As strong as production constraints may be in early phonetic development, infants are sensitive and reactive to the sound patterns around them.

The penetration of an infant's sound repertoire by the ambient language would be facilitated by (a) sensitivity to external auditory patterns, and (b) a mechanism for action–perception linkage, which are both richly supported in the literature. Within the first few months, infants are highly sensitive to the sounds produced in their environment (Chambers, Onishi & Fisher, 2003; Gerken, 2004). A mechanism that links action with perception is the mirror neuron system with Broca's area as its hub (Nishitani, Schurmann, Amunts & Hari, 2005). A reasonable hypothesis is that frequent exposure to a particular sound produced by adults is sufficient to activate sensorimotor systems in an infant's nervous system, presumably in premotor areas such as Broca's area. Therefore, sound patterns in infant vocalizations are determined not only by articulatory constraints or capabilities, but also by the influence of the ambient language, mediated by action–perception linkages.

Production of alveolar affricates, alveolar laterals and palatal affricates by infants may seem surprising, but assuming that infant tongues have muscle fiber composition similar to that in adult tongues, they have a high concentration of fast Type II fibers in the anterior tongue (Stål, Marklund, Thornell, de Paul & Eriksson, 2003), which would facilitate rapid changes in shape and position. Indeed, a high percentage of fast Type II muscle fibers occur in all three muscles critical for producing canonical syllables: anterior tongue, lips and velar-raising muscles. These muscle-fiber characteristics are well suited to executing the rapid articulatory movements presumably inherent to canonical syllables.

Thus, the universal biological substrate of auditory sensitivity, mirror neurons, muscular capability and sensory feedback are interwoven to produce vocalizations in developing humans. This substrate comprises not only constraints but also an adaptive potential that fuses ambient language experience and motor ability into the beginnings of phonetics and phonology.

REFERENCES

Buhr, R. (1980). The emergence of vowels in an infant. *Journal of Speech and Hearing Research* **23**, 73–79.

Carreiras, M. & Price, C. J. (2008). Brain activation for consonants and vowels. *Cerebral Cortex* **18**(7), 1727–35.

Chambers, K. E., Onishi, K. H. & Fisher, C. L. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition* **87**, B69–B77.

Chen, H. P. & Irwin, O. C. (1946). Infant speech: Vowel and consonant types. *Journal of Speech Disorders* **11**, 27–29.

Chen, L.-M. (2005). Development of vocalic production in a Mandarin-learning infant: A longitudinal case study. *Providence Forum: Language and Humanities* **1**(1), 169–92.

Chen, L.-M. & Kent, R. D. (2005). Consonant–vowel co-occurrence patterns in Mandarin-learning infants. *Journal of Child Language* **32**(3), 507–534.

Cheng, C.-C. (1973). *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.

Cheng, C.-M. (1982). Analysis of present-day Mandarin. *Journal of Chinese Linguistics* **10**, 281–358.

Cheng, R. L. (1966). Mandarin phonological structure. *Journal of Linguistics* **2**(2), 135–262.

Davis, B. L. & MacNeilage, P. F. (1990). Acquisition of correct vowel production: A quantitative case study. *Journal of Speech and Hearing Research* **33**, 16–27.

de Boysson-Bardies, B. (1993). Ontogeny of language-specific syllabic productions. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage & J. Morton (eds), *Developmental neurocognition: Speech and face processing in the first year of life*, 353–63. Dordrecht: Kluwer.

de Boysson-Bardies, B., Hallé, P., Sagart, L. & Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language* **16**, 1–17.

de Boysson-Bardies, B. & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language* **67**, 297–319.

de Boysson-Bardies, B., Vihman, M. M., Roug-Hellichius, L., Durand, C., Landberg, I. & Arao, F. (1992). Material evidence of infant selection from target language: A cross-linguistic phonetic study. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (eds), *Phonological development: Models, research, implications*, 369–91. Timonium, MD: York Press.

Gerken, L. A. (2004). Nine-month-olds extract structural principles required for natural language. *Cognition* **93**, B89–B96.

Hsu, J. H. (2003). *A study of the stages of development and acquisition of Mandarin Chinese by children in Taiwan*. Taipei: Crane Publishing Co.

Ingram, D. (1989). *Phonological disability in children*, 2nd edn. London: Whurr Publishers Limited.

Irwin, O. C. (1947*a*). Infant speech: Consonantal sounds according to place of articulation. *Journal of Speech Disorders* **12**, 397–401.

Irwin, O. C. (1947*b*). Infant speech: Consonant sounds according to manner of articulation. *Journal of Speech Disorders* **12**, 402–404.
Irwin, O. C. (1948). Infant speech: Development of vowel sounds. *Journal of Speech Disorders* **13**, 31–34.
Jakobson, R. (1941/68). *Child language, aphasia, and phonological universals*. The Hague: Mouton [English translation, 1968].
Jeng, H.-H. (1979). The acquisition of Chinese phonology in relation to Jakobson's law of irreversible solidarity. *Proceedings of the 9th International Congress of Phonetic Sciences* 2, 155–61. Copenhagen: University of Copenhagen.
Kent, R. D. (1992). The biology of phonological development. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (eds), *Phonological development: Models, research, implications*, 65–90. Timonium, MD: York Press.
Kent, R. D. & Bauer, H. R. (1985). Vocalizations of one-year-olds. *Journal of Child Language* **12**, 491–526.
Kent, R. D. & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America* **72**(2), 353–65.
Kent, R. D. & Vorperian, H. K. (1995). Anatomic development of the craniofacial-oral-laryngeal systems: A review. *Journal of Medical Speech-Language Pathology* **3**, 145–90.
Li, P. J.-K. (1978). Child language acquisition of Mandarin phonology. In R. L. Cheng, Y.-C. Li & T.-C. Tang (eds), *Proceedings of the Symposium on Chinese Linguistics*, 295–316. Taipei: Student Book Co.
Lin, Y. H. (1989). Autosegmental treatment of segmental processes in Chinese phonology. PhD dissertation, University of Texas, Austin.
Liu, I.-M., Chuang, C.-J. & Wang, S.-C. (1975). *Frequency count of 40 000 Chinese words*. Taipei: Lucky Books Company.
Locke, J. L. (1983). *Phonological acquisition and change*. New York: Academic Press.
Luo, C.-C. (1992). Kuo Yü hsüeh [Studies of Mandarin Chinese]. Taipei: Wu-Nan.
MacNeilage, P. F., Davis, B. L. & Matyear, C. L. (1997). Babbling and first words: Phonetic similarities and differences. *Speech Communication* **22**(2–3), 269–77.
Mani, N. & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Development Science* **11**(1), 53–59.
Matyear, C. L. (1998). An acoustical study of vowels in babbling. Unpublished doctoral dissertation, The University of Texas at Austin.
Mowrer, D. E. (1980). Phonological development during the first year of life. In N. J. Lass (ed.), *Speech and language: Advances in basic research and practice*, 99–142. New York: Academic Press.
Nishitani, N., Schurmann, M., Amunts, K. & Hari, R. (2005). Broca's region: From action to language. *Physiology* **20**, 60–69.
Pollock, K. E. (2002). Identification of vowel errors: Methodological issues and preliminary data from the Memphis Vowel Project. In M. J. Ball & F. E. Gibbon (eds), *Vowel disorders*, 83–113. Boston: Butterworth-Heinemann.
Pye, C., Ingram, D. & List, H. (1987). A comparison of initial consonant acquisition in English and Quiché. In K. Nelson & A. van Kleeck (eds), *Children's Language*, Vol. 6, 175–90. Hillsdale, NJ: Erlbaum.
So, L. K. H. & Dodd, B. J. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language* **22**, 473–95.
Stål, P., Marklund, S., Thornell, L. E., de Paul, R. & Eriksson, P.-O. (2003). Composition of human intrinsic tongue muscles. *Cell Tissues Organs* **173**(3), 147–61.
Stoel-Gammon, C. (2002). Intervocalic consonants in the speech of typically developing children: Emergence and early use. *Clinical Linguistics & Phonetics* **16**(3), 155–68.
Stoel-Gamman, C. & Pollock, K. (2008). Vowel development and disorders. In M. J. Ball, M. R. Perkins, N. Muller & S. Howard (eds), *The handbook of clinical linguistics*, 525–48. Malden, MA: Blackwell.
Stokes, S. F. & Wong, I. M. (2002). Vowel and diphthong development in Cantonese-speaking children. *Clinical Linguistics & Phonetics* **16**, 597–617.

Tse, J. K.-P. (1992). Production and perception of syllable final [n] and [ŋ] in Mandarin Chinese: An experimental study. *Studies in English Literature & Linguistics*. Taipei: Department of English, National Taiwan Normal University.

Vihman, M. M. (1992). Early syllables and the construction of phonology. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (eds), *Phonological development: Models, research, implications*, 393–422. Timonium, MD: York Press.

Vihman, M. M., Kay, E., de Boysson-Bardies, B., Durand, C. & Sundberg, U. (1994). External sources of individual differences? A cross-linguistic analysis of the phonetics of mothers' speech to one-year-old children. *Developmental Psychology* **30**, 651–62.

Vihman, M. M., Macken, M. A., Miller, R., Simmons, H. & Miller, J. (1985). From babbling to speech: A reassessment of the continuity issue. *Language* **61**, 397–445.

Wang, F., Huang, H., Chen, C.-W. & Phillips, P. (1986). Articulation development in Mandarin-speaking preschool children ages 3–6 in Taiwan. *Chang Gung Medical Journal* **9**, 68–75.

Waterson, N. (1978). Growth of complexity in phonological development. In N. Waterson & C. E. Snow (eds), *The development of communication*, 415–42. Chichester: Wiley & Sons.

Wong, W. W.-Y. & Stokes, S. F. (2001). Cantonese consonantal development: Towards a nonlinear account. *Journal of Child Language* **28**, 195–212.

Yue-Hashimoto, A. O. (1980). Word play in language acquisition: A Mandarin case. *Journal of Chinese Linguistics* **8**, 181–204.

Zhu, H. (2002). *Phonological development in specific context: Studies of Chinese-speaking children*. Clevedon: Multilingual Matters.

Zhu, H. & Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language* **27**, 3–42.

https://doi.org/10.1017/S0305000909009581 Published online by Cambridge University Press