J. Child Lang. **27** (2000), 3–42. Printed in the United Kingdom

The phonological acquisition of Putonghua (Modern Standard Chinese)*

ZHU HUA AND BARBARA DODD

Department of Speech, University of Newcastle upon Tyne

(*Received* 2 *April* 1998. *Revised* 16 *June* 1999)

ABSTRACT

The phonological acquisition of 129 monolingual Putonghua-speaking children, aged 1;6 to 4;6, is described. Putonghua (Modern Standard Chinese) syllables have four possible elements: tone, syllable-initial consonant, vowel, and syllable-final consonant. The children's errors suggested that Putonghua-speaking children mastered these elements in the following order: tones were acquired first; then syllable-final consonants and vowels; and syllable-initial consonants were acquired last. Phonetic acquisition of the 21 syllable-initial consonants was complete by 3;6 for 75% of children. By 4;6 the children were using the syllable-initial consonants correctly on two thirds of occasions (with the exception of four affricates). Simple vowels emerged early in development. However, triphthongs and diphthongs were prone to systematic errors. Tone errors were rare, perhaps because of their role in distinguishing lexical meaning. In contrast, acquisition of 'weak stress' and 'rhotacized feature' was incomplete in the oldest children assessed. Phonological processes used by the children were identified. Two of

[*] Data collection for the study on which this paper is based was financially supported in part by a University Research Committee small grant from University of Newcastle upon Tyne. Jiang Tao and Shu Hua (Department of Psychology, Beijing Normal University) provided crucial assistance to the data collection. Gerry Docherty (Department of Speech, University of Newcastle) discussed and clarified many phonetics and phonology issues. Li Wei acted as the key consultant for the project as a whole and provided a great deal of input regarding the design of the study and the phonological description of Putonghua. Discussions with him helped to define the notion of 'phonological saliency' more precisely. Transcription reliability was checked by Dom Watt and Li Wei. Alison Holm read and commented in detail on an earlier version of the paper. Parts of the paper have been presented at the 6th International Conference on Chinese Linguistics, 19–21 June 1997, Leiden, the Netherlands and the Child Language Seminar, 4–6 September 1998, Sheffield, UK. Needless to say, the study would not have been possible without the co-operation of the children and their families. Address for correspondence: Zhu Hua, Department of Speech, University of Newcastle upon Tyne, King George VI Building, Queen Victoria Road, Newcastle upon Tyne, NE1 7RU, United Kingdom. e-mail: hua.zhu@newcastle.ac.uk

these processes, syllable-initial consonant deletion and backing, would be considered atypical error patterns in English. Existing theories of phonological acquisition (e.g. concepts of markedness, functional load, feature hierarchies) cannot account for some of the patterns revealed. A satisfactory explanation of the findings requires more attention to the specific characteristics of the linguistic system the children are learning. It is proposed that the saliency of the components in the language system determines the order of acquisition.

INTRODUCTION

Cross-linguistic studies of language acquisition documented in Slobin (1985, 1992, 1995, 1997) pioneered an important area of child language research. This body of work involved: (i) testing the concept of universality of development; (ii) evaluating the applicability of linguistic and psychological theoretical interpretation and explanation; and (iii) differentiating the aspects of development determined by formal linguistic constraints from those determined by social or cognitive factors. Two central components of cross-linguistic phonological acquisition research have emerged: descriptive comparisons of acquisition data; and theoretical interpretations of the cross-linguistic similarities and differences.

Cross-linguistic comparisons of phonological acquisition

The similarities and differences in the developmental patterns of children from various language backgrounds have been examined. The order and rate of acquisition of phonemes and the developmental phonological error processes have been described. Pye, Ingram & List (1987) studied five children learning Quiche, a Mayan language. They found that the children's early phonetic inventories included sounds (e.g. /ʧ, l/) which were not acquired until later by native English-speaking children. Similarly, Jimenez (1987) and Acevedo (1988) found that Mexican–American Spanish-speaking children acquired /t/ and /l/ much earlier than English-speaking children.

Mowrer & Burger (1991) carried out a comparative study of Xhosa- and English-speaking children aged 2;0 to 6;0. They found that Xhosa-speaking children mastered the 20 phonemes shared by Xhosa and English earlier than English-speaking children. The Xhosa children had mastered 31 of the 41 consonants by 3;0, including some affricates (e.g. /ts, ʧ/and clicks. The Xhosa-speaking children also made fewer errors on stops and fricatives than the English-speaking group. However, the two groups were shown to use similar substitution patterns for fricatives, affricates and liquids. The sounds acquired last and most frequently misarticulated by Xhosa-speaking children (e.g. /s, ʃ, r/) were the same phonemes English, German and Swedish children found difficult.

So & Dodd (1995) found that though Cantonese-speaking children's order of consonant acquisition was similar to English-speaking children's, the Cantonese children's acquisition was more rapid. Specific phonological processes used by Cantonese children were also identified. For example, some two-year-olds affricated /s/ (e.g. [pa tsi] for /pa si/; [tsoej] for /soej/). This pattern would be unusual in English-speaking children who acquire affricates later than fricatives. The more common developmental error for English-speaking children involves stopping of affricates (e.g. [tɪp] for /tʃɪp/). In contrast, while some Cantonese two-year-olds deaffricated /ts/ (e.g. [siw] for /tsiw/), affrication of /s/ was much more common.

These cross-linguistic comparisons of phonological acquisition have raised two important questions: why do children living in totally different linguistic and social environments acquire the same set of phonemes at nearly the same stage? When some phonemes are common to two languages, why does their order of acquisition differ and why do their error patterns vary?

Phonological acquisition theories

Theories of phonological acquisition need to account for evidence from cross-linguistic studies. Two major issues need to be addressed. One issue is the universal tendencies in children's phonological acquisition; the other is the role language-specific features play in determining the phonological development of the children of a given language.

Jakobson (1941/1968) suggested that whether a sound would be acquired early could be explained by the distribution of the sound amongst the world's languages. According to his 'laws of irreversible solidarity', nasals, front consonants and stops (found in virtually all the languages) would be acquired earlier than their oppositions, orals, back consonants and fricatives respectively. He proposed that there were certain sounds which were more basic and central to all human languages and these sounds would therefore would be acquired earlier than other sounds. Jakobson's view of phonological acquisition in terms of oppositions or contrasts set the agenda for the subsequent studies of child phonology.

The notion of 'markedness' has been used to interpret the similarities and differences in the order of sound acquisition (Eckman, 1977; Anderson, 1983). It was hypothesised that those sounds which appear early in a child's inventory are maximally unmarked, while those occurring late are marked. Therefore, children would use unmarked sounds as substitutions for marked sounds. Edwards' (1974) study of English-speaking children aged 1;8 to 3;11 found that children usually substituted the unmarked member for those marked contrasts (e.g. [s] for /ʃ/), but details varied from one child to another and from one developmental stage to another.

Some researchers found that the traditional labels in the taxonomy of oppositions such as voice, place and manner of articulation were not adequate

when explaining the order of acquisition of phonemes. A more detailed descriptive unit was adopted: the feature. The feature system focused on the articulatory differences between phonemes (see Chomsky & Halle, 1968). Among the most important features were those distinguishing between vowels and consonants (sonorant, vocalic, consonantal); those distinguishing the sounds in terms of place of articulation (anterior, coronal, high, low, back and rounded); and those distinguishing the sounds in terms of manner of articulation (nasal, lateral, continuant, delayed release and strident). Each phoneme was a combination of several features. It was hypothesized that unmarked features would be acquired first because unmarked features are considered more phonetically natural. Therefore, children would tend to replace marked features with unmarked features. A number of studies (e.g. Irwin & Wong, 1983; Yavaş, 1997) have applied the feature system to the analysis of children's speech.

Dinnsen (1992) proposed that there might be a universal hierarchical structure with a highly limited set of ordered features applicable to the phonetic inventories of all languages. Each feature in the hierarchy has a number of default specifications (i.e. unmarked values). Children's acquisition would therefore be a process of replacing a default value with a language-specific value. The phoneme acquisition order of a particular language corresponds with the dominance relationships and default values: features ranked highly in the hierarchy would be acquired early; default features would be acquired before non-default features. Dinnsen's model offers an alternative account for cross-linguistic similarities and differences in the order of phoneme acquisition. However, the explanatory power of his model has so far rarely been tested with the phonological acquisition of children other than English- and Spanish-speaking children. Determining the set of default and non-default values common to all languages is a continuing goal of phonological research.

Jakobson's 'laws of irreversible solidarity' and later feature theory sought to explain children's acquisition of sounds in the structure of the language they learned and emphasized the innate nature of acquisition. In contrast, other researchers (Locke, 1980, 1983; Kent, 1992) emphasized the role of articulatory and perceptual constraints on children's acquisition of phonology. Locke (1980) argued that far from simply being a physical process, 'perception is very much constrained by one's sense of phonological structure and lexical expectation'. Developmental speech forms can therefore be accounted for by the perceptual similarity between sounds that occur frequently in children's babbling and sounds that occur infrequently. Locke (1983) proposed three universal mechanisms of development: maintenance; learning; and loss. When a child passes the babbling stage and starts to acquire a target phonological system, certain sounds are maintained from the babbling repertoire. Sounds not present in the babbling repertoire are then

developed through interactions in the linguistic environment (a learning process). The child must also relinquish and lose the 'extrasystemic sounds', sounds existing in the babbling repertoire but not in the target phonological system. The interaction of these three mechanisms results in the acquisition of the target phonology.

Pye *et al.*'s (1987) study of Quiche-speaking children challenged Locke's theory. They attributed differences between the phonological acquisition of Quiche and English to the specific nature of the two phonological systems. Pye *et al.* argued that articulatory and perceptual constraints could not account for the earlier acquisition of /l/ and /ʧ/ by Quiche-speaking children than English-speaking children. They explained the differences they found by introducing the concept of 'functional load' which was first proposed by the Prague School. Functional load refers to the relative importance of each phoneme within a specific phonological system. However, how to calculate it is still a matter of controversy. Pye *et al.* determined the functional load of a phoneme by its frequency of occurrence in oppositions or minimal pairs. For example, /l/ and /ʧ/ are acquired earlier by Quiche children because these sounds carry a greater functional load in the Quiche phonological system than in English.

Functional load is, however, difficult to measure (Catford, 1988). Pye *et al.* (1987) admitted that phonemes with a high frequency of occurrence might not necessarily carry a high functional load. For example, /ð/ occurs in a small class of frequent words (such as *the*, *this*, etc.) and is thus the second most frequent fricative in English. However, the functional load of /ð/ is quite small, since 'we could change all English /ð/ into [d]s and still communicate' (Ingram, 1989). Despite this, Pye *et al.* measured the functional load of syllable-initial consonants in Quiche by counting the frequency of syllable-initial consonants occurring in the 500 most commonly used words of five- and six-year-old children.

There are two problems with Pye *et al.*'s method of determining functional load:

(1) There is no guarantee that sounds frequently used by children are significant for a phonological system;
(2) The rank-order of frequencies for syllable-initial consonants common to Quiche and English did not support the similarities and differences found in the children's order of acquisition. For example, the sound /w/ was ranked as the second most frequently used in Quiche and as seventh in English, indicating that /w/ should be acquired earlier in Quiche than in English. In fact, it was acquired at the same age in both languages.

So & Dodd (1995) were also critical of Pye *et al.*'s measurement of functional load: other aspects of phonology that may contribute to the

functional load of consonants, such as vowel, syllable structure, and tone had not been considered. Despite these weaknesses, the notion of functional load does explore the relationship between the order of phoneme acquisition and the role of these phonemes in a given language environment. Theory needs to account for language specific phenomena in acquisition.

Children's phonological acquisition is a highly complex process and influenced by a variety of sources. It is conceivable that none of the theories discussed so far account for both the universal tendencies and the language-specific patterns that have been found. Further cross-linguistic research on children's phonological acquisition is needed, focusing on both the identification of universal tendencies and the influence of the ambient language.

THE STUDY

This paper describes the phonological acquisition of Putonghua (Modern Standard Chinese, or MSC). Despite the fact that the Chinese language has the largest native-speaking population in the world, there have been very few studies of language acquisition of Chinese children. The so-called Chinese language is in fact a variety of languages or 'dialects', most of which are mutually unintelligible. For example, one of the major 'dialectal groups' of the Chinese language is *Beifanghua*, 'Northern speech', spoken by over 70% of the population in China. In English, Beifanghua is often referred to as Mandarin. Beifanghua is very different in all structural aspects, especially in phonology, from Cantonese, another major 'dialectal group'. Cantonese is spoken widely in the south of China, Hong Kong and in the Chinese immigrant communities in the English-speaking world. Putonghua, literally 'common speech', is the language promoted by the government throughout China since 1950. It is based on the phonological and grammatical system of Beifanghua and is widely used in the mass media and taught in schools. It is a variety of Mandarin that has to some extent been standardized.[1] The following description is based on Chao (1968) and Norman (1988).

THE PHONOLOGICAL SYSTEM OF PUTONGHUA

Putonghua syllable has four possible structures: V, CV, VC and CVC. There are 21 consonants which can occur in syllable-initial position and two (i.e. /n/ and ŋ/) which can occur in syllable-final position.

[1] Although as a standardized language variety the phonological system of Putonghua is based on that of northern Chinese dialects (represented most notably by Beijing Mandarin), there are significant differences between Putonghua and Beijing Mandarin concerning prosody, tones, weak stress, and rhotacization. Some of these differences are discussed in Norman (1988) and Lin (1990).

Consonants

The place and manner of articulation of Putonghua consonants are described in Table 1. Aspiration serves as a distinctive feature of Putonghua consonants.

TABLE 1. *Place and manner of articulation of Putonghua consonants*

	Bilabial	Labio-dental	Alveolar	Retroflex	Alveolo-palatal	Velar
Stop	p p^h		t t^h			k k^h
Nasal	m		n			ŋ
Affricate			ts ts^h	tʂ $tʂ^h$	tɕ $tɕ^h$	
Fricative		f	s	ʂ	ɕ	x
Approximant				ɹ		
Lateral approximant				l		

There are six pairs of aspirated and unaspirated consonants and all of them are voiceless. There are three alveolo-palatal phonemes, which seldom occur in other languages (Ladefoged & Maddieson, 1996).

Vowels

Vowels can be classified into three groups with 9 simple vowels, 9 diphthongs and 4 triphthongs. The nine simple vowels are /i, y, u, ɤ, o, ʌ, ə, ɛ, ɚ/. Among these vowels, /ɚ/ is a retroflexed central vowel. It occurs either in isolation or in rhotacization and thus has very restricted combination with onset consonants. /ə/ and /ɛ/ have very restricted use. /ə/,as a monophthong, occurs only in weakly stressed syllables (see below). /ɛ/ is used only in conversational particles, which express a speaker's emotion, such as surprise, agreement. /i/ has three allophones (Norman, 1988; termed as 'fricative vowels' in Ladefoged & Maddieson, 1996):

- occurring after /ts/, /ts^h/ and /s/, it represents a weak syllabic prolongation of the preceding consonants, (usually transcribed as /ɿ/ in the traditional Chinese phonetic transcription);
- occurring after retroflexes /tʂ/, /$tʂ^h$/, /ʂ/ and /ɹ/, it represents a weak syllabic retroflex continuant, (usually transcribed as /ʅ/ in the traditional Chinese phonetic transcription);
- occurring after all other consonants, it represents a high front unrounded /i/.

The diphthongs can be divided further into offglides and onglides: /ae/, /ei/, /ɑo/ and /ow/ are offglides, the first vowel sound being longer and

https://doi.org/10.1017/S030500099900402X Published online by Cambridge University Press

having more intensity; /iA/, /iɛ/, /uA/, /uo/, and /yɛ/ are onglides, the second element being sonorous. In all of the four triphthongs, the middle vowel has the most intensity and is the longest.

Tone and tone sandhi

There are four tones in Putonghua, i.e. high level, high rising, falling-rising and high falling, primarily characterized by voice pitch but also by length and intensity. They are referred to as tones 1, 2, 3 and 4 respectively in this paper. Differences in tones can change the meaning of a lexical item.

Tone sandhi in Putonghua is closely associated with the morphological structure of Chinese words, and falls into the following three categories:

- The falling-rising tone will become a rising tone before another fall-rising tone, if they are in one meaning group (e.g. disyllabic words or phrases); before any other tone, it will only retain the falling part of its contour. In effect, it is only in isolation or before a pause (usually at word or phrase boundaries) that the falling-rising tone completes its full contour.
- The high falling tone, followed by another high falling tone, will become low falling tone.
- There are four lexical items which follow their own rules. They are /pu4/ ('no'), /i1/ ('one'), /tɕi1/ ('seven') and /pa1/ ('eight'). /pu4/, /tɕil/ and /pa1/ will become high rising tone before a high falling tone. /i1/ will become a high falling tone before all the tones except a high falling tone and become a high rising tone before a high falling tone.

Weak stress

Weak stress is an essential prosodic feature of Putonghua, which, like tone sandhi, is associated with Chinese morphology. Weakly stressed syllables have a very short duration and a much reduced pitch range. The pitch of a weakly stressed syllable is primarily determined by the preceding tone: it is half-low after a high level tone; mid after a high rising tone; half-high after a falling-rising tone; and low after a high falling tone (for a fuller discussion, see Norman, 1988). Other phonetic changes associated with weak syllables include unaspirated voiceless consonants being fully voiced, and vowels being centralized. For example, /pitsi/ (meaning *nose*) is realized as [pidzə] or [pidzɐ], with the pitch of the syllable /tsi/ changing from falling-rising tone to mid level. Such changes are lexically determined and not restricted to grammatical suffixes and particles. There are three typical categories:

- 'suffix' type in which a single syllable noun is combined with a weakly stressed suffix 'zi' (/tsi/); for example, /pitsi/ (nose), /tɕʰyntsi/ (skirt).

- 'reduplication' type in which the reduplicated second syllables are weakly stressed; for example, /ɕiŋɕiŋ/ (star) and /ɕiɛɕiɛ/ (thank you).
- 'lexeme' type in which two lexemes, especially nouns, are combined together and the second lexeme is generally weakly stressed, for example, /ɚtuo/ (ear), /t^{h}owfA/ (hair).

Rhotacization

Putonghua has a number of rhotacized syllables. With regard to semantics, rhotacization often indicates that the referent is something common, familiar or small. In some cases, it may carry either a diminutive or a slightly pejorative implication. Though morphemically the rhotacization is a feature attached to a syllable at syllable-final position (e.g. /xuAɹ/, flower), the rhotacization process affects both the vowels and consonants in the coda. Changes that may co-occur include: the syllable-final consonant /n/ may be dropped; the syllable-final consonant /ŋ/ will be dropped but the vowel will become nasalized; a front vowel may become centralized. Though theoretically all syllables can be rhotacized, whether a syllable is rhotacized in speech and how frequently a syllable is rhotacized is determined lexically. Some words are always rhotacized and not using the rhotacized form of the word would be unusual, e.g. /nyxae/ (girl), /yantɕ^{h}yan/ (circle). Some are optionally rhotacized, e.g. /mən/ (gate), /tɕ^{h}itʂ^{h}ɤ/ (car).

There are controversies over some of the issues concerning the description of Putonghua phonological system. For example, the sound that /ɹ/ represents is sometimes described as /ʐ/ and the sound that /x/ represents is described as /χ/ (e.g. Norman, 1988). Ladefoged & Maddieson (1996) argued that retroflex should be labelled as flat post-alveolar while alveolo-palatal should be palatalized post-alveolar. A detailed phonetic analysis is beyond the scope of this paper. We use the labels and symbols which are widely used in the literature in our discussion of Putonghua. Differences between the phonology of Putonghua and that of English are summarized in Table 2.

The status of /i/, /u/ and /y/ in triphthongs and on-gliding diphthongs is another issue of controversy. They have been traditionally labelled as 'medial' or 'prenucleus glide' between the syllable-initial consonant and the following vowel, while recent studies (e.g. Wang, 1989) suggested they should be considered as part of onset rather than part of coda. The implications of these arguments for interpreting developmental patterns of phonological acquisition of Putonghua children will be discussed in the Discussion section.

TABLE 2. *Differences between Putonghua and English phonology*

	Putonghua	English
Tones	4 tones	None
Syllable-initial consonants	p, p^h, t, t^h, k, k^h m, n f, s, ɕ, x, ʂ l, ɹ ts, ts^h, tɕ, $tɕ^h$, tʂ, $tʂ^h$	p, b, t, d, k, g m, n θ, ð, f, v, s, z, ʃ, ʒ, h l, ɹ ʧ, ʤ
Clusters	None	p b t d k g θ+l r j s+p t k l w s+p t k+l r j w
Syllable-final consonants	n, ŋ	m, n, ŋ p b t d k g θ, ð, f, v, s, z, ʒ, ʃ, ʧ, ʤ l, ɹ
Vowels	i, u, y, o, ɤ, ʌ, ə, ɛ, ɚ ae, ei, ɑo, ow, iʌ, uo, yɛ iɑo, iow, uae, uei,	i, ɪ, ɛ, æ, ʌ, a, ɒ, ɔ, ʊ, u, ɜ, ə eɪ, əʊ, aɪ, au, ɔɪ, ɪə, ɛə, ɔə, ʊə (eɪə, əʊə, aʊə, ɔɪə)
Syllable structure	$[C_{0-1}]$ -V- $[C_{0-1}]$	$[C_{0-3}]$ -V- $[C_{0-4}]$

PHONOLOGICAL ACQUISITION OF CHINESE-SPEAKING CHILDREN

The phonological acquisition of Chinese children remains under-explored. Erbaugh's (1992) detailed review of the acquisition of Mandarin which is largely based on studies carried out in Taiwan, for example, did not describe aspects of phonology, except for noting error-free tonal acquisition. Lee's (1996) more recent literature review listed only three case studies that described the phonological acquisition of Mandarin-speaking children: Li (1977); Jeng (1979); and Shiu (1990). Lee highlighted the heavy influence of Jakobson's earlier work on these child phonology studies: they all tested the validity of Jakobson's laws of irreversible solidarity. For example, Jakobson claimed that the acquisition of front consonants preceded back consonants. Li's (1977) data on Taiwan Mandarin revealed that palatal and velar consonants were acquired early, and the velar nasal often replaced the dental nasal at syllable-final position. (Note: Li described /n/ and /ŋ/at syllable-final position as dental and velar nasals, which was different from the oft-cited descriptions of Mandarin). The limited information available suggests that Chinese children's phonological development is influenced by the structure of the language and therefore is different from that of children of other linguistic backgrounds (Chao, 1951/1973; Li, 1977; Li & Thompson, 1977; Clumeck, 1977; Jeng, 1979; Shiu, 1990). However, the previous Chinese phonological acquisition research is restricted to diary records of one or two

subjects. Such data do not allow generalizations concerning order or age of phoneme acquisition, nor identification of the developmental error patterns used by Chinese-speaking children.

This paper reports a normative cross-sectional study of 129 Putonghua-speaking children aged 1;6 to 4;6. The order of phoneme acquisition, age of phoneme acquisition, and phonological processes used by these children are identified. The language-specific features of Putonghua indicate that developmental patterns would reflect an interaction of universal tendencies and language-specific constraints. Specific hypotheses regarding the individual aspects of the phonology were formed:

Order of phoneme acquisition

(a) Nasals would be acquired earlier than orals, and stops earlier than fricatives.This pattern, predicted by Jakobson (1941/1968), is already supported by the existing studies of Cantonese-, Japanese-, Italian-, Turkish-, Spanish- and English-speaking children (for Cantonese data, see So & Dodd, 1995; for Japanese, see Yasuda (1970), cited in Locke, 1983; for Italian, see Bortolini & Leonard, 1991; for Turkish, see Topbas, 1997; for Spanish, see Jimenez, 1987; Acevedo, 1988; for English, see Olmsted, 1971; Prather, Hedrick & Kern, 1975).

(b) Marked features, aspiration and affrication, would be acquired later than the default and unmarked features of the language (for aspiration, see So & Dodd, 1995; for affrication, see Olmsted, 1971; Prather *et al.*, 1975.)

Phonological processes

Simplification processes would be evident both in the overall structure of syllables and within syllable components. Young children were expected to deaspirate and stop within a syllable, and assimilate and delete some syllable components such as syllable-initial and -final consonants at syllable level.

Tones

The acquisition of tone would be early, probably due to its capacity in differentiating lexical meaning and fulfilling children's communicative intentions.

Vowels

Children would make fewer vowel errors compared with syllable-initial consonant errors, probably because vowels are compulsory components of syllables in Putonghua.

METHODS

Subjects

The phonological acquisition of 134 children aged 1;6 to 4;6 was assessed. The children were recruited from five nurseries and kindergartens in Beijing. School records and parent reports ensured that all of the children were acquiring Putonghua as their first language, and had no intellectual or hearing impairment. Analysis of the data revealed that five children had very atypical speech errors and were therefore excluded from the present normative data analysis. The 129 children were divided into six age groups of six month intervals. The subject characteristics of each group are outlined in Table 3.

TABLE 3. *Number of subjects in each age group*

Age group (mean age)	Girl	Boy	Total
1;6–2;0 (1;10)	10	11	21
2;1–2;6 (2;3)	11	13	24
2;7–3;0 (2;9)	11	10	21
3;1–3;6 (3;4)	13	13	26
3;7–4;0 (3;9)	11	15	26
4;1–4;6 (4;3)	5	6	11
Sum	61	68	129

Materials

Picture-naming and picture-description tasks were used. For the picture-naming task, 44 words were used to sample all of the tones and all of the phonemes in each legal word position (see Appendix 1). The word list included 39 nouns likely to be known by young children (e.g. nose, apple, bird, bed, sun). Four verb phrases (thank you, bye-bye, brush teeth, wash face) frequently used with young children in daily interactions were included in the picture-naming task. One colour term was also used (red). High quality colour drawings that were attractive to children were prepared. The drawings were laminated on A5 white cards. Five pictures of scenes incorporating most of the objects and actions in the picture-naming task were also prepared. Due to the lack of information on the frequency distribution of Putonghua phonemes in speech, the choice of target words and phrases was primarily motivated by their familiarity to young children and imagibility for producing the pictures. the frequency of phonemes in the test varied. Appendix 2 summarizes frequency distribution of syllable-initial consonants, vowels, syllable-final consonants, tones, tone sandhi, weak stress and rhotacization in the picture-naming task.

Procedure

The children were assessed individually in a quiet room at their nursery or kindergarten by a trained examiner. If the child failed to say the target word in the picture naming task, the examiner would offer semantic or contextual prompts. If it was impossible to elicit a spontaneous production of the target word, the child would be asked to imitate the examiner. Imitated responses were noted on the record form. If the child produced the target word incorrectly, the examiner would ask the child to say the word again. Up to three attempts were made to elicit the correct pronunciation of the word. Repetitions were noted on the record form.

The children were also asked to describe what they could see in the five picture scenes. To elicit spontaneous continuous speech, the children were asked either 'Can you tell me what's happening here?' or 'What's funny about this picture?'. Each session lasted between 10 and 15 minutes and was audiotaped using a Sony professional micro recorder.

Imitated production

Although we tried to use the commonest words and phrases in the picture-naming task, our objective of sampling all the phonemes in each legal word position and tones meant that some of the words and phrases used may be less familiar to some children than others. There were occasions when the child failed to produce the target word spontaneously, for the following reasons:

(a) The target word was beyond the child's conceptual and lexical ability and therefore the task of accessing an appropriate lexical representation distracted the child from fulfilling the phonological task.

(b) The child was actively using avoidance strategy when s/he found certain sounds too difficult to pronounce (Macken & Ferguson, 1983).

In the present study, the children were asked to imitate the examiner when they failed to produce the target word or phrase spontaneously. Imitated responses were taken into account only when age of phoneme emergence was calculated, since the focus was on children's articulation ability. When age of phoneme stabilization (i.e. articulation accuracy) was calculated imitated responses were excluded. It should be noted that the frequency of occurrence of imitated responses decreased with age. The mean frequency and standard deviation (in parentheses) of occurrence of imitated responses for each age group in our sample was: 1;6–2;0: 12·04 (6·01); 2;1–2;6: 6·79 (4·11); 2;7–3;0: 4·42 (2·92); 3;1–3;6: 2·73 (2·61); 3;7–4;0: 1·27 (1·48); 4;1–4;6: 0·81 (0·88). Statistical analysis showed that there was a significant age effect in the frequency of occurrence of imitation responses (one-way ANOVA: $F = 28{\cdot}7599$, $p = 0{\cdot}0000$).

There were some words which children, especially those of the younger age groups, tended to fail to produce spontaneously, due to conceptual,

lexical or cultural reasons. The five most frequently imitated words were /ny χae/ (girl, 25·6 %), /ɕin/ (heart, 25·6 %), /tɕʰyn tsi/ (skirt, 21·7 %), /nan χae/ (boy, 20·9 %), and /ɕi liɛn/ (wash face, 16·2 %). The five least frequently imitated words were /tɕʰiow/ (ball, 0·02 %), /ɕiɑo tɕʰi tʂʰʏ/ (car, 0·02 %), /ɕi kuA/ (watermelon, 0·04 %), /ɚ tuo/ (ear, 0·04 %), and /ʂuA iA/ (brush teeth, 0·05 %).

Transcription

The data from the picture-naming and picture description tasks were transcribed by a phonetician who is a native speaker of Putonghua and experienced in transcribing children's speech. Incomplete responses (use of a shortened word, similar to the use of 'plane' for 'airplane' in English) were recorded and excluded in the data analysis. Recordings from 20 subjects were independently transcribed by another phonetician to check the transcription reliability. The inter-transcriber reliability for syllable-initial word-initial consonants, syllable-initial within-word, syllable-final word-final, syllable-final within-word consonants and vowels was 97·6 %, 94·6 %, 98·1 %, 97·8 % and 98·5 % respectively.

DATA ANALYSIS

Comparison of connected and single word speech

Many researchers assume that children would make more errors in continuous speech than in single words because connected speech is linguistically more complex (e.g. constructing sentences) (Shriberg, Austin, Lewis, McSweeny & Wilson, 1997). It is necessary to test this assumption to determine the speech mode most indicative of phonological ability. The production of all of the words elicited in both speech conditions by 18 children (3 from each age group) was compared.

Consistency of production

The variation in production of the same phoneme in the same syllable position was investigated. The variation may be an indication of children's development when there is variation between the correct target and a developmental error. In contrast, a range of error forms used for the same phoneme in the same phonetic context may reflect acquisition difficulties (Dodd, 1995). Consistency in the realization forms of the same sound in picture naming and spontaneous speech was compared to investigate the stability of the children's realizations.

Phoneme emergence

A phoneme was considered to have emerged when 90 % of the children in an age group produced the sound at least once, irrespective of whether it was the

correct target. This measure determined when the children were able to articulate each sound.

Phoneme stabilization

Each phoneme occurred in the sample once or several times. Since there is certain amount of inconsistency in children's production, a criterion is needed to derive age of phoneme stabilization. A sound was considered stable when the child produced the sound correctly on at least two of three opportunities. When 90% of the children in an age group achieved an accuracy rating of at least 66·7% (i.e. 2/3) for a phoneme, the phoneme would be considered to have been stabilized by that age group. To balance various actual occurrence of phonemes in the task, To determine the proportion of correct productions of each phone from the number of opportunities in the sample the following accuracy rating formula was applied (see Shriberg & Kwiatkowski, 1982; Shriberg, *et al.*, 1997):

$$\text{Percentage accuracy} = \frac{\text{the number of times of a phoneme produced correctly}}{\text{the number of opportunities for a phoneme}}$$

Since phoneme development is a continuum ranging from the initial stage of being able to articulate a sound in isolation to the final stage of being able to articulate a sound both phonetically and phonologically accurately, it is important to define the terms and criteria in describing phoneme acquisition. Following previous studies (Prather *et al.*, 1975; So & Dodd, 1995), this study adopted a 90% criterion in determining age of acquisition. Data on 75% of the age groups are also presented for comparison with other studies. One of the reasons for setting up a 90% criterion is that prevalence figure for phonologically delayed and disordered children is reported to be about 10% of the normal population (National Institute on Deafness and Other Communication Disorders, 1994).

Phonological processes

The consistent differences between children's realizations and adult's target forms are described as phonological processes. In the data, if 10% of the children of the same age group were found to use the same or similar phonemes (in terms of place or manner of articulation) to replace certain target sounds, that process would be recorded. Percentage of the children in each age group using that process was calculated to measure the developmental pattern of the process involved.

RESULTS

Comparison of connected and single word speech

The children's speech accuracy in continuous speech was comparable to their accuracy in the picture-naming task. Of the 18 children, seven gave more correct responses in picture-description than they did in picture-naming; five gave the same number of correct responses in both modes; and six gave more correct responses in picture-naming (see Table 4). A *t*-test ($p = 0{\cdot}714$) indicated there was no significant difference between the two speech samples in terms of the number of correct responses and errors.

TABLE 4. *Comparison of speech production in picture-naming and picture-description tasks* (*based on 18 children randomly selected from all age groups*)

Age	Sex	Number of syllables targeted in both picture-naming and -description	Number of correct syllables in picture-naming	Number of correct syllables in picture-description
2;0	F	24	9	11
2;0	M	9	2	2
2;0	M	12	7	7
2;3	M	30	13	13
2;3	M	34	21	22
2;4	F	26	15	18
2;8	M	20	12	18
2;8	F	21	13	10
2;9	F	29	16	17
3;1	M	30	25	21
3;2	F	26	17	19
3;3	M	44	38	38
3;9	F	29	28	26
3;9	F	41	27	29
3;10	M	31	29	29
4;1	F	41	38	34
4;1	M	30	28	27
4;2	F	43	42	41

Consistency of production

The children's realizations of sounds in different linguistic contexts were highly consistent. Individual sound productions were compared between the connected speech and single word tasks. Vowel consistency was 98·9%; syllable-initial word-initial consonant consistency was 84·4%; syllable-initial within-word consonant consistency was 95·3%; syllable-final word-final consonant consistency was 87·7%; and syllable-final within-word consonant consistency was 96·9%.

Overview of speech errors

Errors were classified into three types: syllable-initial; syllable-final; and vowel errors. The mean number of all types of errors decreased with age (see Table 5). Compared to the mean error of 30·76 in the youngest age group, the

TABLE 5. *Overview of speech error in different age groups*

Age group	1;6–2;0	2;1–2;6	2;7–3;0	3;1–3;6	3;7–4;0	4;1–4;6
Total error	646	625	347	354	229	82
Mean error	30·76	26·04	16·52	13·61	8·81	7·45
Syllable-initial error (%)	65·8	69·6	72·9	70·1	73·8	79·3
Syllable-final error (%)	14·2	12·8	10·7	17·8	17·0	9·8
Vowel error (%)	20·0	17·6	16·4	12·1	9·2	10·9

oldest group's mean error was only 7·45. The proportion of syllable-initial consonant errors in the total number of errors was greater than that of vowel and syllable-final consonant errors. The proportion of syllable-initial error increased with age range while the proportion of vowel error decreased, suggesting vowel acquisition was complete earlier than that of syllable-initial consonants.

Emergence of syllable-initial consonants

By 4;6, 90% of the children were able to articulate all the 21 syllable-initial consonants (Table 6). Among the first sounds produced by 90% of the

TABLE 6. *Age of emergence of syllable-initial consonants*

	90% criterion	75% criterion
1;6–2;0	t, t^h, k, m, n, x, tɕ, $tɕ^h$, ɕ	t, th, k, m, n, f, s, x, tɕ, $tɕ^h$, ɕ, p^h, p
2;1–2;6	f, s, tʂ	ʂ, tʂ, $tʂ^h$, k^h
2;7–3;0	p, l	ts, l
3;1–3;6	p^h, k^h, $tʂ^h$	ɹ, ts^h
3;7–4;0	ʂ	
4;1–4;6	ts, ts^h, ɹ	
> 4;6		

children were nasals; alveolar stops; alveolo-palatal fricatives and affricates; and the velar stop and fricative. The two alveolar affricates and the alveolar approximant appeared later. Zn terms of features, some unaspirated sounds emerged earlier than their aspirated pairs (e.g. /k/ and /k^h/); some unaspirated sounds emerged more or less simultaneously with aspirated pairs (e.g. /ts/ and /ts^h). The six affricates occurred later than the stop /t/ which

has the same place of the articulation. However, the continuants at the same place of articulation did not necessarily appear before the affricates. For example, /ts/ and /ts^{h}/ emerged later than /s/, but /tʂ/ and /tʂ^{h}/ emerged earlier than /ʂ/. The three alveolo-palatals emerged very early in the children's speech.

Stabilization of syllable-initial consonants

Age of stabilization of syllable-initial consonants are summarized in Table 7. stabilization of phonemes in Putonghua phonology, compared with their age of emergence, can be categorized into the following three groups:

- phonemes which were stabilized as soon as the child was able to utter them. These phonemes were basically error free, e.g. /t/, /m/, /p/;
- phonemes which took a relatively short period to become stabilized after the child was able to utter them, e.g. /n/, /f/, /x/; and
- phonemes which took a long time to become stabilizzed after the child was able to utter them, e.g. /tɕ/, /tɕ^{h}/, /s/.

TABLE 7. *Age of stabilization of syllable-initial consonants*

	90% criterion	75% criterion
1;6–2;0	t, m	t, t^{h}, m, n, x
2;1–2;6	n	p, p^{h}, k, k^{h}, ɕ, tɕ, tɕ^{h}
2;7–3;0	p, t^{h}, f, x, ɕ	f
3;1–3;6	k, k^{h}	—
3;7–4;0	p^{h}	—
4;1–4;6	l, s, ɹ, tɕ, tɕ^{h}	l, s, ʂ, ɹ
> 4;6	ʂ, tʂ, tʂ^{h}, ts, ts^{h}	tʂ, tʂ^{h}, ts, ts^{h}

Vowels

Vowels emerged very early in development. The youngest group of children were able to pronounce all of the simple vowels. Vowel errors were classified into the following categories (see Table 8):

- Triphthong reduction: triphthongs were often reduced to diphthongs (in most cases) or sometimes to monophthongs. The middle vowel, the main vowel in Putonghua triphthongs, was maintained and one of the other vowels was deleted. This error pattern was most evident for the triphthong /iɑo/: 37% of the children reduced this triphthong. Of these children, 29% used /iA/ and only 8% used /ɑo/. The second most frequently reduced triphthong was /uei/ (10% of the children): the most frequent reduction form was /ei/ (7% of the total subjects).

https://doi.org/10.1017/S030500099900402X Published online by Cambridge University Press

TABLE 8. *Percentage of children using processes affecting vowels in different age groups* (%)

Age group	Triphthong reduction	Diphthong reduction	Substitutes	Assimilation
1;6–2;0	67	67	71	10
2;1–2;6	63	58	83	4
2;7–3;0	48	38	76	0
3;1–3;6	23	19	50	0
3;7–4;0	23	19	42	0
4;1–4;6	22	9	45	0

- Diphthong reduction: diphthongs were often reduced to a simple vowel. The vowel retained was the louder and more sonorant vowel of a diphthong. The children tended to produce the second element of ongliding diphthongs when reduced. For example, 12% of the children realized /ua/ as /a/ once or several times in their speech production and none of them realized it as /u/. The first element of offgliding diphthongs was most often maintained. Thus, more children replaced /ɑo/with [ʌ] than with [o].
- Vowel substitution: some of the children substituted vowels at the same time when they deleted consonants (34·7% of vowel substitutions). When a syllable-final consonant was deleted, the vowel was lengthened, most frequently with the vowel /ɛ/. As a result, a monophthong in the target syllable would turn into a diphthong ending with [ɛ], a diphthong would turn into a triphthong ending with [ɛ]. Other vowel substitution errors were not systematic.

Tone

Tone errors were rare, even in the youngest group of children. Only two tone errors were observed in the entire data corpus and they were produced by the children of the youngest age group. The two tone errors were [ɕiɛ4] for /ɕiɛ2/ and [wan2] for /wan3/. Five children in the youngest age group and three children in the second youngest age group occasionally used citation tones when tones should be adjusted according to tone sandhi rules. However, due to the nature of the cross-sectional study and the type of picture-naming task used in data collection, we prefer to defer any conclusions on the acquisition of tone sandhi by Putonghua-speaking children. As Li & Thompson (1977) pointed out, a child who is able to adjust tones in single word context may not necessarily have acquired the tone sandhi rule. It is likely that s/he manages to learn the single words as adjusted forms without being aware of tone sandhi rules. An analogy is that an English-speaking child who uses 'went' may not know anything about the past tense. Alternatively, the

scarcity of tone sandhi errors in the study may be an artifact of the cross-sectional design, in that tone sandhi rules may be acquired during a very short period of time and such a study is unable to capture such rapid changes. We are presently undertaking a longitudinal study of four children aged 1;0 to 2;0 which may offer us further insights on tonal acquisition.

Stress

Data concerning the acquisition of weakly stressed syllables are shown in Table 9. Weakly stressed syllables were evident in 57% of the youngest

TABLE 9. *Percentage of children using weak stress in different age groups* (%)

Age group	Emerge*	50% correct	60% correct	70% correct	80% correct	90% correct	100% correct
1;6–2;0	57	5	0	0	0	0	0
2;1–2;6	88	8	4	4	0	0	0
2;7–3;0	100	57	38	29	10	5	0
3;1–3;6	100	77	69	62	19	4	0
3;7–4;0	100	96	88	77	58	38	15
4;1–4;6	100	100	100	91	73	45	36

* Compared to the cases when the children could stress 50%–100% of the target weak stresses correctly, 'emerge' refers to the case when the children were able to produce the target weak stress once or several times.

group of children's speech. Only 36% of the oldest group were able to correctly stress all the 13 weak stresses. Almost all of the weak stress errors were associated with the pitch level and duration. Of the 'suffix' weak syllable type, 42·3% of the total errors were weakly stressed syllable deletions. For example, [pi2] for /pi2 tsi0/. The remaining errors involved pitch level and intensity: the children either used citation tone of /tsi3/, i.e. the falling-rising tone, or lengthened the syllable.. Of the other two weak stress syllable types, most of the errors (93·4%) occurred when children used the citation tone (the pitch level of the syllable when pronounced individually, e.g. /tow2fA0/ was realized as [tow2fA4]).

Rhotacized feature

Only 57 % of the youngest group used the rhotacized form once or several times (see Table 10). However, over 90% of the children over 2;0 and all the children over 3;6 used rhotacized forms. The rhotacized feature was not acquired all at once. With four words which are always rhotacized in adult speech, some children rhotacized some of them while using non-rhotacized forms for the others.

TABLE 10. *Percentage of children rhotacizing target words in different age groups* (%)

Age group	Emerge*	uan	mən	ɕiŋɕiŋ	xuA	niɑo	ɕywnmɑo	tɕitʂʰɤ	nyxae	nanxae	yantɕʰyan
1;6–2;0	57	10	14	14	14	10	5	0	43	19	33
2;1–2;6	92	13	29	25	25	4	4	0	67	83	42
2;7–3;0	95	10	24	29	43	14	14	0	71	76	52
3;1–3;6	96	19	38	42	42	12	4	4	85	85	62
3;7–4;0	100	19	42	62	85	19	31	4	96	92	62
4;1–4;6	100	18	27	45	73	18	18	9	100	91	91
Total	90	15	30	36	46	12	12	10	76	74	54

* 'Emerge' refers to the case when the children were able to produce rhotacization once or several times.

TABLE 11. *Phonological processes affecting syllable-initial consonants and percentage of children using these phonological processes in all age groups*

	Percentage of children using processes (%)							Most common error types and percentages of children in all age groups using error types
	1;6–4;6	1;6–2;0	2;1–2;6	2;7–3;0	3;1–-3;6	3;7–4;0	4;1–4;6	
Assimilation	43	86	42	52	39	8	11	IC of one syllable harmonizes with IC of another syllable. 21% e.g. ɕiaŋ. tɕ iɑo: dɑŋ. tɑo IC harmonizes with final consonant by being nasalized. 17% e.g. ɕi.liɛn: ɕi.niɛn; yɛ.liɑŋ: yɛ.liɑŋ: yɛ.niɑŋ/myɛ.niɑŋ IC harmonizes with final velar by being velarized. 2% e.g. tsuɑŋ: xuɑŋ
IC deletion	37	81	58	29	15	15	33	IC deletion before high vowels /i/, /y/ & /u/. 37% e.g. tsuei: uei ; tɕʰyn: yn /l/ deletion. 16% e.g. liɛn;liɑŋ; iɑŋ
Fronting	87	91	100	91	85	81	89	Retroflex fricatives and affricates become alveolars. 77% e.g. ʂ; s; tʂ: ts; tʂʰ: tsʰ Alveolo-palatal fricatives & affricates become post-alveolars; 36% e.g. ɕ: ʃ; tɕ: ʧ Velar stops become alveolar stops; 16% e.g. k: t; k: d; kʰ: tʰ
Backing	70	91	96	67	58	54	44	Alveolar affricates and fricative become post-alveolar; 65% e.g. ts: ʧ/dʒ; s: ʃ Fricatives become a glottal fricative; 5% e.g. f: h; ʂ: h; x: h
X-velarization	48	76	71	48	31	27	22	X-velarization occurs before a high vowel /u/; 24%. e.g. ʂu: xu X-velarization occurs before a high vowel /i/ & [y]; 23%. e.g. ɕi: xi /f/ becomes X-velarized; 8%. e.g. fA: xA; fei: xei

TABLE 11 (*cont.*)

	Percentage of children using processes (%)							Most common error types and percentages of children in all age groups using error types
	1;6–4;6	1;6–2;0	2;1–2;6	2;7–3;0	3;1–-3;6	3;7–4;0	4;1–4;6	
Stopping	63	95	92	76	50	46	11	tɕ/tɕʰ/tʂ/tʂʰ/ts/tsʰ: t/d/tʰ; 63% x:k/g; 22% ɕ/n/l/ʂ: t/d; 13%
Affrication	34	67	29	24	27	4	0	s/ʂ/ɕ: ʧ/ ʤ; 34% ɕ: tɕ/tɕʰ; 22%
Deaspiration	56	48	88	67	50	46	44	tʂʰ: ts; 21% tɕʰ: ʧ; 20% tʂʰ/tsʰ/tʰ: t/d; 16% tɕʰ: tɕ; 8% kʰ: k; 5% pʰ: p; 2%
Aspiration	32	24	46	43	27	12	0	t : tʰ; 15% ɕ/tɕ: tɕʰ; 15% k: kʰ; 3%
Gliding	28	43	33	57	23	12	0	ɹ: [j]; 28%

Phonological processes

Phonological processes affecting syllable-initial position. The phonological processes affecting syllable-initial position can be generalized into three groups: assimilation, deletion and systematic substitution. Table 11 summarizes the data. The most typical realization of these processes and their sub-categories are outlined below:

- Assimilation. Assimilation occurs when one or more distinctive features of a sound is transferred to an adjacent sound. The transference can take place both within a syllable and across syllables and is thus highly context-sensitive. Twenty-one per cent of the children harmonized a syllable-initial consonant with another consonant and 17% of the children nasalized syllable-initial consonants. Both progressive and regressive assimilation were found in the data.
- Deletion. Syllable-initial consonant deletion was very common in the youngest group. It happened most frequently before the vowels /i/, /y/ and /u/. For example, a number of the children deleted /l/ in the target word /liɛn/ and /ts/ in /tsuei/.
- Fronting. While the most typical fronting pattern is the realization of target velar sounds as alveolars in English-speaking children, only 16% of the Putonghua-speaking children in this study have used this pattern. The majority of the children (77%) fronted the retroflex sounds by realizing them as alveolars and 36% replaced the alveolo-palatals with post-alveolars, which do not exist in Putonghua phonology.
- Backing. Backing occurs when the place of articulation is backed. This category is rarely reported in the studies of other languages. However, in terms of percentage of the subjects using the process, it is the second most frequent process used by Putonghua-speaking children: 65% of the children substituted post-alveolars for alveolars. For example, /suA/ was realized as [ʃuA].
- X-velarization. X-velarization was another frequent form of backing, and so frequent that it has been categorized as a group of its own for clarity: 48% of the children used [x], a velar fricative, to replace other fricatives and affricates. In most cases, X-velarization process occurred either before the vowel /u/ or before the vowels /i/ or /y/.
- Stopping and affrication. The most common type of stopping (63%) in the data was the use of stops of the same place or nearest place of articulation in the place of affricates. Its opposite process, affrication, was used by a relatively small number of the children (34%).
- Deaspiration and aspiration. Deaspiration (56%) occurred significantly more frequently than the aspiration process (32%) and was

often associated with other processes such as deaffrication and fronting. Among all the aspirated sounds, the aspirated retroflex /tʂʰ/ and alveolo-palatal /tɕʰ/ were most frequently deaspirated while /pʰ/ was rarely deaspirated.

- Gliding. /ɹ/was replaced with [j] by 28 % of the children. Besides this type of substitution, there were 4 % of the children replacing /ɹ/with the liquid [l].

Phonological processes affecting syllable-final position. In Putonghua, there are only two possible syllable-final consonants and they are both nasal, /n/ and /ŋ. Both of these syllable-final consonants occurred very early in the children's inventory. In the data, all the children in the youngest age group were able to articulate these two phonemes. The five phonological processes associated with these two syllable-final consonants were /n/ deletion, /ŋ/ deletion, replacing /n/ with [ŋ], replacing /ŋ/ with [n] and syllable-final consonant addition. Examples are listed in Table 12.

TABLE 12. *Phonological processes affecting syllable-final consonants and percentage of children using these phonological processes in all age groups*

	Targets	Examples	Percentage of children
/n/ deletion	san	sa	57
/ŋ/ deletion	pʰiŋ	pʰi	29
Replacing /n/ with [ŋ]	san	saŋ	55
Replacing /ŋ/ with [n]	pʰiŋ	pʰin	3
Syllable-final consonant addition	niɑo	niɑŋ	6

The two most frequently used processes were /n/ deletion and replacing /n/ with [ŋ], used by 57 % and 55 % of children respectively. /ŋ/ deletion was ranked the third most common process. However, the other two processes, syllable-final consonant addition and replacing /ŋ/ with [n] rarely occurred. It is worth noting that all the syllable-final consonants in the children's speech were 'legal' nasal consonants. No other consonant occurred in syllable-final position even among the youngest age group. These processes were unidirectional in the sense that deletion of syllable-final consonants could occur but not addition of syllable-final consonants; replacement of /n/ with [ŋ] could occur, but not the reverse.

DISCUSSION

The phonological acquisition of 129 monolingual Putonghua-speaking children, aged 1;6 to 4;6, is described. Putonghua syllables have four possible

elements: tone, syllable-initial consonant, vowel, and syllable-final consonant. The children's errors suggested that Putonghua-speaking children acquired these elements in the following order: tones were acquired first; then syllable-final consonants and vowels; and syllable-initial consonants were acquired last. Phonetic acquisition of the 21 syllable-initial consonants was complete by 3;6 for 75% of children. By 4;6 the children were using the syllable-initial consonants correctly on two thirds of occasions (with the exception of four affricates). Simple vowels emerged early in development. However, triphthongs and diphthongs were prone to systematic errors. Tone errors were rare. In contrast, acquisition of 'weak stress' and 'rhotacized feature' was incomplete in the oldest children assessed. Compared with children acquiring other languages, Putonghua-speaking children generally shared the tendency of structural and systemic simplifications in their production. There were also some differences in the processes used by the children acquiring Putonghua. For example, syllable-initial consonant deletion and backing, which are considered atypical error patterns in English, were evident in the speech of the children acquiring Putonghua.

Phoneme acquisition

The order of phoneme acquisition in Putonghua provides evidence for and against various theories of acquisition. Jakobson's (1941/1968) law of irreversible solidarity predicts that nasals should be acquired before orals, front consonants before back consonants, and stops before fricatives. The Putonghua-speaking children acquired nasals before orals, and stops before fricatives. However, front consonants (/p/, /p^h/, /m/ and /f/) were acquired at about the same stage as back consonants (/k/, /k^h/, /x/ and /ŋ/). The three alveolo-palatal sounds, which are very rare in the world major languages, were acquired relatively early (75% of children by 2;6). These data do not support Jakobson's proposal that the frequency of a phoneme across the world's languages determines the order of acquisition of the phoneme.

The last 10 phonemes to be acquired in Putonghua include all the three retroflex sounds, all the six affricates and both liquids. The late acquisition of these sounds, believed to be difficult to articulate and perceive (Locke, 1983), supports the hypothesis that biological constraints affect the order of phoneme acquisition. However, a closer comparison of the age of *emergence* of syllable-initial consonants and the age of *stabilization* of these consonants (see Tables 6 and 7) casts some doubt on the role of articulatory constraints. Some of the later-stabilized sounds *emerged* very early in children's speech. In the youngest age group, 90% of the subjects were able to articulate the affricates /tɕ/ and /$tɕ^h$/ once or more. By 2;6, 90% of children were able to utter the retroflex /tʂ/ and alveolar fricative /s/. However, it was not until

the children were aged over 4;0 that they began to consistently use these sounds correctly. The delay between emergence and stabilization indicates that articulatory constraints were not a major factor in the phonological acquisition.

Feature acquisition

There was a clear developmental sequence to feature acquisition. The features of aspiration, affrication and retroflex were acquired last. Late acquisition of affrication has been reported for a variety of languages. English-speaking children acquired the two affricates in English (i.e./ʧ/ and /ʤ/ later than other phonemes (Olmsted, 1971; Prather *et al.*, 1975). So & Dodd (1995) also found that Cantonese-speaking children acquired the two affricates after all the other phonemes. Timm (1977, cited in Locke, 1983) claimed that in Russian the affricate /ʧ/ ranked 8th of the 33 consonants in terms of error scores.

However, it is premature to conclude that affrication is a marked feature. Locke (1983) discussed the research into other languages that have a different pattern of affrication acquisition. For example, two affricates were among the first group of phonemes to be acquired in Japanese (Yasuda, 1970, cited in Locke, 1983). Battacchi, Facchini, Manfredi & Rubatta (1964, cited in Locke, 1983) also reported the early acquisition of the affricate /ʧ/ by Italian children.

The discrepancies associated with the acquisition of a particular feature such as affrication highlight the possibility of the influence of the ambient language on acquisition. The cross-linguistic differences also reflect the explanatory inadequacies of the theoretical concept of markedness or default features. The current theories are able to explain acquisition order similarities. However, they do not account for cross-linguistic differences.

Phonological processes

Phonological processes are 'a mental operation that applies in speech to substitute, for a class of sounds or sound sequences presenting a common difficulty to the speech capacity of the individual, an alternative class identical but lacking in the difficult property' (Stampe, 1973: 1). Stampe viewed processes as an indicator of restricted resources (motor-oral skills, cognitive capacity, perceptive ability) available to children at a particular stage of development. Ingram (1986) proposed a more active role for the children acquiring their phonological system. He viewed phonological processes as 'a universal set of hierarchically ordered procedures used by children to simplify speech' (Ingram, 1986: 223–224). Therefore, these processes can be interpreted as children's simplification strategies, if we

TABLE 13. *Phonological processes used by more than 10% of Putonghua-, Cantonese-, English-, or Italian-speaking children of different age groups*

	Putonghua		Cantonese (So & Dodd, 1995)		English (Grunwell, 1982)		Italian (Bortolini & Leonard, 1991)	
1;6–2;0	Assimilation	Fronting ʂ → s	—	—	Reduplication	Fronting of velars	—	—
	IC deletion	Backing ts → tʃ	—	—	Consonant	Stopping	—	—
	Triphthong reduction	X-velarization	—	—	harmony	Gliding r → w	—	—
	Diphthong reduction	Stopping ts → t	—	—	Final consonant	Context sensitive	—	—
	Final n deletion	Affrication	—	—	deletion	voicing	—	—
	Final ŋ deletion	Deaspiration	—	—	Cluster reduction	—	—	—
	n → ŋ	Aspiration	—	—	—	—	—	—
	—	Gliding ɹ → j	—	—	—	—	—	—
2;1–2;6	Assimilation	Fronting ʂ → s	Cluster reduction	Stopping f/s/ts/ts^{h}	Final consonant	(Fronting of velars)	Assimilation	Liquid deviation
	IC deletion	Backing ts → tʃ	Assimilation	Fronting k^{h} → t	deletion	Stopping	Weak syllable	Obstruent
	Triphthong reduction	X-velarization	/h/ deletion	Deaspiration	Cluster reduction	Gliding r → w	deletion	devoicing
	Diphthong reduction	Stopping ts → t	—	Affrication	—	Context sensitive	Cluster reduction	Spirantization
	Final n deletion	Affrication	—	—	—	voicing	Metathesis	—
	Final deletion	Deaspiration	—	—	—	—	Epenthesis	—
	n → ŋ	Aspiration	—	—	—	—	—	—
		Gliding ɹ → j	—	—	—	—	—	—
2;7–3;0	Assimilation	Fronting ʂ → s	Cluster reduction	Stopping f /s/ts/ts^{h}	—	Stopping /v ð z tʃ ʤ/	—	—
	IC deletion	Backing ts → tʃ	Assimilation	Fronting k^{h} → t	—	ɵ → f	—	—
	Triphthong reduction	X-velarization	—	Deaspiration	—	Fronting ʃ → s	—	—
	Diphthong reduction	Stopping ts → t	—	Affrication	—	Gliding r → w	—	—
	Final n deletion	Affrication	—	Deaffrication	—	Context sensitive voicing	—	—
	Final ŋ deletion	Deaspiration	—	—	—	—	—	—
	n → ŋ	Aspiration	—	—	—	—	—	—
	—	Gliding ɹ → j	—	—	—	—	—	—

TABLE 13 (*cont.*)

	Putonghua		Cantonese (So & Dodd, 1995)		English (Grunwell, 1982)		Italian (Bortolini & Leonard, 1991)	
3;1–3;6	Assimilation	Fronting ʂ → s	Cluster reduction	Stopping f /s/ts/ts^h	—	Stopping /v ·/	—	—
	IC deletion	Backing ts → ʧ	—	Fronting k^h → t	—	ɵ → f	—	—
	Triphthong reduction	X-velarization	—	Deaspiration	—	Fronting of /ʧ ʤ ʃ/	—	—
	Diphthong reduction	Stopping ts → t	—	—	—	Gliding r → w	—	—
	Final n deletion	Affrication	—	—	—	—	—	—
	Final ŋ deletion	Deaspiration	—	—	—	—	—	—
	n → ŋ	Aspiration	—	—	—	—	—	—
	—	Gliding ɹ → j	—	—	—	—	—	—
3;7–4;0	IC deletion	Fronting ʂ → s	Cluster reduction	Stopping	—	ɵ→ f	—	—
	Triphthong reduction	Backing ts → ʧ	—	—	—	ð → d/v	—	—
	Diphthong reduction	X-velarization	—	—	—	Palatalization of /ʧ ʤ ʃ/	—	—
	Final n deletion	Stopping ts → t	—	—	—	Gliding r → w	—	—
	n → ŋ	Deaspiration	—	—	—	—	—	—
	—	Aspiration	—	—	—	—	—	—
	—	Gliding ɹ → j	—	—	—	—	—	—
4;1–4;6	IC deletion	Fronting ʂ → s	—	—	—	—	—	—
	Triphthong reduction	Backing ts → ʧ	—	—	—	—	—	—
	Diphthong reduction	X-velarization	—	—	—	—	—	—
	Final n deletion	Stopping ts → t	—	—	—	—	—	—
	n → ŋ	Deaspiration	—	—	—	—	—	—

assume that the children's underlying representations resemble those of the adult[2]. This inherent tendency to simplify underlies children's attempts to realize the target pronunciation, irrespective of language. Consequently, all children are expected to have common processes of structural and systemic simplification.

The phonological processes used by the children acquiring Putonghua phonological system distinctly revealed both universal tendencies and language-specific constraints on acquisition. Structural simplifications such as asassimilation, deletion, and reduction and systemic substitutions such as stopping, fronting, backing, gliding were evident in Putonghua-speaking children's speech sample. These processes are similar across languages. Table 13 compares the phonological processes used by Putonghua-, Cantonese-, English-, and Italian-speaking children. Despite discrepancies in terminology and analysis method, the structural simplification processes are very similar. There are noticeable cross-linguistic differences, however, in the systemic substitution processes. Some of these differences can be attributed to the language-specific phonological characteristics.

The six aspirated/unaspirated phoneme pairs of Putonghua allowed exploration of the acquisition process. While deaspiration and aspiration processes were both evident, deaspiration was more prevalent (56% of the children deaspirated). Although the children continued to deaspirate phonemes throughout the age groups described, aspiration was suppressed earlier. Cantonese-speaking children's use of aspiration and deaspiration processes is similar to Putonghua children's (So & Dodd, 1995). These patterns suggest that children acquire the unmarked before the marked member of a pair irrespective of language. These findings support the existence of universal tendencies in cross-linguistic phonological acquisition.

Some error patterns, such as syllable-initial consonant deletion, indicate the children's high sensitivity to the characteristics of Putonghua phonology. Syllable-initial consonant deletion always occurred before the vowels /i/, /u/ or /y/. This pattern may reflect the flexible function of these three vowels. The vowels /i/, /u/ and /y/ can occur either as monophthongs, or as a component of diphthongs and triphthongs, such as /iɛ/ or /uei/. As mentioned before, the status of these vowels has been debated. Traditionally these sounds have been described as the first element of a diphthong or triphthong: they function as 'medial' or 'prenucleus glide' between the

[2] Since it is difficult to determine to what extent a child's underlying representation resembles that of an adult, the phonological processes – the consistent differences between children's realizations and target forms – may alternatively be interpreted as a realization of children's own phonological rules and part of their own phonological systems. In this regard, when a syllable deletion process takes place, the child may produce the simplified form without being aware of the full target form.

syllable-initial consonant and the following vowel, having a shorter duration than the following main vowels. Several studies (e.g. Wang, 1989) have challenged this traditional description and proposed that the Putonghua syllable has a branching onset consisting of a consonant and a glide (i.e. syllable-initial consonant and /i/, /u/ or /y/). In other words, these three vowels form part of syllable-initial clusters. Therefore, the process of syllable-initial consonant deletion could be considered as cluster reduction.

Acquisition of the liquid /ɹ/ has been widely discussed in cross-linguistic studies. Bortolini & Leonard (1991) compared the acquisition of this phoneme phone in several languages. They discovered that /ɹ/ was frequently replaced by [l] in Italian-, Hindi-, Igbo-, Portuguese-, Quiche- and Spanish-speaking children. However, English-speaking children substituted the glide [w]. In this study, 28% of Putonghua-speaking children substituted [j], another glide, for /ɹ/. Only 4% of the children replaced the sound with [l]. Bortolini & Leonard argued that the cross-linguistic difference between English and Italian was due to the restricted use of /w/ in Italian, as well as phonetic differences in each language. However, similar restrictions do not account for Putonghua-speaking children's use of /j/ rather than /w/. In Putonghua, /w/ and /j/ are variants of medial or pre-nucleus vowels /u/ and /i/. They are equally flexible in their combinations with other vowels or consonants. Other factors must affect the pattern of children's systemic simplification of speech.

Factors affecting systemic simplification

The acquisition of Putonghua phonology has shown that children do not simplify their speech by replacing difficult sounds with sounds that are easier to articulate. The long delay between emergence and stabilization of some phonemes (particularly the three alveolo-palatals) undermines the role of biological constraints on stage of acquisition. Further, while markedness may account for some error patterns (e.g. the unidirectional replacement of aspirated sounds with unaspirated sounds), it does not account for cross-linguistic differences in the acquisition order of affricates.

The concept of functional load directly links order of phoneme acquisition to the role of those phonemes in a language's phonological system. However, previous proponents of functional load (e.g. Pye *et al.*, 1987) have failed to investigate the influence of aspects of phonology other than consonants on order of acquisition (So & Dodd, 1995). A simplified analysis of the impact of functional load, measuring only the load of consonants, does not explain the acquisition order of Putonghua. For example, Pye *et al.* (1987) argued that /ŋ/ has lower functional load in English than /m/ because /ŋ/ does not occur initially and thus has a smaller number of oppositions. In the same vein, /ŋ/ should also have a lower functional load in Putonghua

than /n/, since /ŋ/ does not occur word-initially. Consequently, children should acquire /n/ before /ŋ/. However, as shown in Table 13, the children made more errors on /n/. They either deleted /n/ or substituted /ŋ/. The phoneme /ŋ/, with a lower functional load, was acquired before /n/. These data suggest that Pye *et al.*'s (1987) notion of functional load is inadequate. Alternatively, the order of acquisition might be determined by the phonological saliency of a component within the language.

Phonological saliency

The notion of phonological saliency has been alluded to by others (e.g. Peters, 1983; Vihman, 1996), but there is no agreement on its definition. In the context of the current study, we use phonological saliency as a syllable-based, language-specific concept. It is determined and affected by a combination of several factors:

(a) The status of a component in the syllable structure, especially whether it is compulsory or optional; a compulsory component is more salient than an optional one;
(b) The capacity of a component in differentiating lexical meaning of a syllable; a component which is more capable of distinguishing lexical information is more salient than one which carries less lexical information;
(c) The number of permissible choices within a component in the syllable structure. e.g. 21 syllable-initial consonants would be considered less salient compared to 4 tone contrasts.

Tone has the highest saliency in Putonghua: it is compulsory for every syllable; change of tone would change lexical meaning; and there are only four alternative choices. Lexical information of a word in Putonghua is conveyed by both tone and phoneme sequence. Therefore, tone is crucial in differentiating lexical meaning. In contrast, other syllable components are less vital: information lost by an incorrect phoneme within a phoneme sequence can be remedied to some extent by other phonemes in the sequence (e.g. in English we could guess that [lelou] means *yellow*). The phoneme sequence as a whole unit shares the task of conveying lexical meaning. Therefore, the significance of each phoneme in a sequence is less than each tone.

Tone is more salient than the three other syllable components for a variety of reasons, not simply because it conveys meaning. Syllable-initial consonants have the lowest saliency of the four syllable components: their presence is optional (not all syllables have syllable-initial consonants); and there is a range of 21 syllable-initial phonemes that can be used. Vowels are compulsory

syllable components. However, the relatively large number of options (21 in total including monophthongs, diphthongs and triphthongs) lowers their saliency. Although there are only two syllable-final consonants, their saliency is undermined by their optional presence in the syllable structure. Compared with the saliency of tones, consonants and vowels, the saliency of 'weak stress' and 'rhotacized feature' is much lower. Neither 'weak stress' nor 'rhotacized feature' are compulsory syllable components, and their value in differentiating lexical meaning is low.

The saliency values of the four syllable components in Putonghua are congruent with their acquisition order: tones were acquired earlier than syllable-final consonants and vowels, which were acquired earlier than syllable-initial consonants. The features of ' weak stress' and rhotacization were acquired last due to their low saliency value.

Differences in the saliency of individual components in different languages may result in the cross-linguistic variations in developmental patterns. The number of options within a syllable component may determine rate of acquisition when other factors are equal. For example, Putonghua-speaking children's tonal acquisition was more rapid than that of Cantonese-speaking children. In both languages tone is a compulsory syllable component and differentiates lexical meaning. However, Cantonese has nine tones while Putonghua has only four. The four children in So & Dodd's (1995) longitudinal study mastered only three of the nine tones by 1;6, and their acquisition was not complete until 2;0. In the present study, only two of the children aged 1;6 to 4;6 made any tone errors.

The effect of the number of options within a component was also evident in the number of vowel errors made by Cantonese and Putonghua-speaking children. The Putonghua-speaking children made more vowel errors than Cantonese-speaking children of the same age. Putonghua's 21 vowels include monophthongs, diphthongs, and triphthongs. In contrast, Cantonese has 22 vowels, but only two vowel types: monophthongs and diphthongs. The additional vowel type in Putonghua reduces the saliency of the vowel component, resulting in the slower acquisition of vowels.

The role of phonological saliency in determining acquisition rate is compatible with previous research findings. So & Dodd (1995) reported that the consonant acquisition rate of Cantonese-speaking children was more rapid compared to the acquisition rate of English-speaking children. Cantonese-speaking children acquired their range of consonants by 3;6. English-speaking children's phoneme repertoires were not complete until they were five-years-old (Prather *et al.*, 1975). Mowrer & Burger (1991) found that Xhosa-speaking children acquired most consonant phonemes earlier than their English-speaking counterparts. These discrepancies in consonant acquisition rates between Cantonese, English and Xhosa are compatible with the concept of saliency (see Appendix 3 for comparison of the phonological

structure of Putonghua, Cantonese, English and Xhosa). Cantonese has only 17 consonants and 2 clusters, while English has 24 consonants and 49 clusters. Although consonants are optional syllable components in both languages, the larger number of consonants and clusters in English lowers the saliency of each consonant. Therefore, the acquisition rate of English is slower than Cantonese.

It is important to remember that phonological saliency as we define here is a language-specific concept. The saliency level of a particular phonological feature is determined by its role within the phonological system of the language, not by reference to other languages. For instance, although Xhosa has 41 consonants, it has a very simple syllable structure. A typical Xhosa syllable is structured as CV. In addition, Xhosa has very few consonant clusters.[3] The relatively indispensable status in a syllable and lack of clusters thus contribute to the higher saliency of consonants in Xhosa and explain their early acquisition. Nevertheless, the factors that should be taken into account in determining the saliency of a phonological component require further testing of hypotheses with cross-linguistic studies.

It should also be noted that the concept of phonological saliency is different from linguistic markedness in that it is cognitive in nature and characterizes the accessibility or noticeability of certain linguistic forms to children. Slobin (1979, 1985) also discussed the notion of saliency with particular regards to the acquisition of grammatical structure. He proposed a series of ' operating principles', by which the learner perceives and processes the linguistic input and organizes it in his / her internal system.

Interaction between lexical and phonological acquisition

The acquisition of weak stress and rhotacization in Putonghua reflects interactions between lexical and phonological development. The children's tendency to use citation tones in the place of weak stress may be attributed to caretakers' often exaggerated and emphasized manner of speaking in which citation tones are given to weak stress (Li & Thompson, 1977; Erbaugh, 1992). However, it is arguable whether the acquisition of these features are rule-based or lexically motivated. As our data show, most of the children in the study made consistent errors with the weak stress syllable /tsi/ when it occurred in different syllable contexts. If the learning of weak stress took place on a word-by- word basis, different error types would have been present. Further investigation is needed to examine the interaction of phonological and lexical constraints in children's phonological and lexical development.

[3] Lanham (1969) suggested that the so called 'nasal compounds (i.e. cluster)' should be best treated as 'prenasalized consonants', i.e. single segments.

CONCLUSION

The present study describes the acquisition of phonology of 129 Putonghua-speaking children aged 1;6 to 4;6. Typical developmental patterns of Putonghua-speaking children are discussed. It is important to bear in mind that the present study is a cross-sectional one. Our conclusions about age of acquisition should be interpreted in that context. We are currently undertaking further research into phonological development and change over time using longitudinal data. Furthermore, acquisition of certain aspects of phonology, such as tone and tone sandhi, awaits further evidence from longitudinal studies.

Given that no reliable information is available on the frequency distribution of Putonghua phonemes in natural speech, our selection of words and phrases in preparing the stimuli for data collection was primarily driven by their familiarity to young children and imagibility for producing the pictures. Consequently, the frequency of phonemes varied (see Appendix 2 for details). This may have affected the results to some extent. However, as the phonemes which had the same frequency in the picture-naming task showed different age of emergence and stabilization, it was unlikely that the overall finding was an artifact of the seemingly unbalanced frequency distribution.

Despite the methodological limitations, Tthe developmental patterns in terms of order of phoneme acquisition and phonological processes identified in Putonghua-speaking children have a number of implications for the theoretical interpretation of cross-linguistic similarities and differences which have been reported in the existing literature:

- Cross-linguistic differences in the order of phoneme acquisition cannot be accounted for in terms of the frequency of the phonemes across the world's languages;
- Nor can they be explained by appealing to the biological constraints or articulatory limitations of young children;
- Although there is a clear developmental sequence in terms of 'feature', the theoretical concept of universal 'markedness' or 'default features' has a number of explanatory inadequacies;
- There are language-specific influences on the order of phoneme acquisition. However, while the current proposal of 'functional load' directly links the order of phoneme acquisition to the role of these phonemes in a language, it fails to investigate the impact of aspects of phonology other than consonants on the order of acquisition; moreover, there are difficulties in the measurement of functional load;
- Some cross-linguistic variations in the rate and order of acquisition of vowels, consonants and prosodic features such as tone are better

accounted for by the concept of saliency. Components with higher phonological saliency would be acquired earlier than components with lower saliency. There is a need for refining the notion of 'phonological saliency' that might be able to capture cross-linguistic differences in phonological acquisition and development.

REFERENCES

Acevedo, M. A. (1988). *Development of Spanish consonants in three to five year olds*. Paper presented at the Annual ASHA Convention, Boston.

Anderson, J. L. (1983). The markedness differential hypothesis and syllable structure difficulty. In G. S. Nathan (ed.), *Proceedings of the Conference on the Uses of Phonology*. Carbondale: Southern Illinois University.

Battacchi, M. W., Facchini, G. M., Manfredi, M. M. & Rubatta, C. O. (1964). Presentazione di un reattivo per l'esame dell'articolazione fonetica nei fanciulli in eta prescolare di lingua italiana. *Bollettino della Societa Italiana di Fonetica, Foniatria e Audiologia* **13**, 441–86.

Bortolini, U. & Leonard, L. B. (1991). The speech of phonologically disordered children acquiring Italian. *Clinical Linguistics & Phonetics* **5**, 1–12.

Catford, J. C. (1988). Functional load and diachronic phonology. In Y. Tobin (ed.), *The Prague School and its legacy*. Amsterdam: John Benjamins Publishing Company.

Chao, Y. R. (1951/1973). The Cantian idiolect: an analysis of the Chinese spoken by a twenty-eight-month-old child. In C. A. Ferguson & D. I. Slobin (eds), *Studies of child language development*. New York: Holt, Rinehart and Winston, INC.

—— (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Clumeck, H. (1977). *Studies in the acquisition of Mandarin phonology*. Unpublished doctoral dissertation, University of California, Berkeley.

Dinnsen, D. (1992). Variation in developing and fully developed phonetics inventories. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (eds), *Phonological development: models, research, implications*. Timonium, MD: York Press.

Dodd, B. (1995). *Differential diagnosis & treatment of children with speech disorder*. London: Whurr Publishers.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning* **27**, 315–30.

Edwards, M. L. (1974). Perception and production in child phonology: the testing of four hypotheses. *Journal of Child language* **1**, 205–19.

Erbaugh, M. S. (1992). The acquisition of Mandarin. In D. Slobin (ed.), *The crosslinguistic study of language acquisition*. Vol. 3. Hillsdale, NJ: Erlbaum.

Ingram, D. (1986). Phonological development: production. In P. Fletcher & M. Garman (eds), *Language acquisition*. Cambridge: C.U.P.

—— (1989). *First language acquisition: method, description and explanation*. Cambridge: C.U.P.

Irwin, J. V. & Wong, S. P. (eds). (1983). *Phonological development in children 18 to 72 months*. Carbondale: Southern Illinois University Press.

Jakobson, R. (1941/1968). *Child language, aphasia and phonological universal*. The Hague: Mouton.

Jeng, Heng-hsiung. (1979). The acquisition of Chinese phonology in relation to Jakobson's laws of irreversible solidarity. In *Proceedings of the 9th International Congress of Phonetic Sciences*. University of Copenhagen.

Jimenez, B. C. (1987). Acquisition of Spanish consonants in children of aged three to five years seven months. *Language Speech and Hearing Services in Schools* **18**, 357–61.

Kent, R. (1992). The biology of phonological development. In C. A. Ferguson, L. Menn & C. Stoel-Gammon (eds), *Phonological development: models, research, implications*. Timonium, Maryland: York Press.

Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Lanham, L. W. (1969). Generative phonology and the analysis of Nguni consonants. In *Lingua* **24**, 155–62.
Lee, Thomas H.-T. (1996). Theoretical issues in language development and Chinese child language. In C.-T. James Huang & Y.-H. Audrey Li (eds), *New horizons in Chinese linguistics*. Dordrecht: Kluwer Academic Publishers.
Li, C. N. & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language* **4**, 185–99.
Li, Paul J.-K. (1977). Child language acquisition of Mandarin phonology. In R. Cheng, Y. C. Li & Ting-chi Tang (eds), *Proceedings of the Symposium on Chinese Linguistics : 1977 Linguistic Institute of the Linguistic Society of America*. Taipei: Student Books.
Lin, T. (1990). *Yuyin tansuo jigao*. Beijing: Beijing Language Institute.
Locke, J. (1980). The prediction of child speech errors: implications for a theory of acquisition. In G. H. Yeni-komshian, J. F. Kavanagh & C. A. Ferguson (eds), *Child phonology*, 1: *production*. New York: Academic Press.
Locke, J. (1983). *Phonological acquisition and change*. New York: Academic Press.
Macken, M. & Ferguson, C. A. (1983). Cognitive aspects of phonological development. In K. E. Nelson (eds), *Children's language*. Vol. 4. Hillsdale, NJ: Erlbaum.
Mowrer, D. & Burger, S. (1991). A comparative analysis of the phonological acquisition of consonants in the speech of two and a half and six year old Xhosa- and English-speaking children. *Clinical Linguistics and Phonetics* **5**, 139–64.
National Institute on Deafness and Other Communication Disorders (1994). *National strategic research plan*. Bethesda, MD: Department of Health and Human Services.
Norman, J. (1988). *Chinese*. Cambridge: C.U.P.
Olmsted, D. (1971). *Out of the mouths of babes*. The Hague: Mouton.
Peters. A. (1983). *The units of language acquisition*. Cambridge: C.U.P.
Prather, E., Hedrick, D. & Kern, C. (1975). Articulation development between two and four years. *Journal of Speech and Hearing Disorders* **40**, 179–91.
Pye, C., Ingram, D. & List, H. (1987). A comparison of initial and final consonant acquisition in English and Quiche. In K. E. Nelson & A. Van Kleek (eds), *Children's language*. Vol. 6. Hillsdale, NJ: Erlbaum.
Shiu, Huei-shiou (1990). *The phonological acquisition by Mandarin-speaking children : a longitudinal case study on children from 9 months through three years old*. Unpublished MA thesis, Taiwan Normal University.
Shriberg, L. D. & Kwiatkowski, J. (1982). Phonological disorders III: a procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders* **47**, 256–70.
—— , Austin, D., Lewis, B. A., McSweeny, J. L. & Wilson, D. L. (1997). The Percentage of Consonants Correct (PPC) Metric: extensions and reliability data. *Journal of Speech, Language, and Hearing Research* **40**, 708–22.
Slobin, D. (1979). *Psycholinguistics*. Second edition. Glenview, ILL.: Scott, Foresman and Company.
—— (ed). (1985, 1992, 1995, 1997). *The crosslinguistic study of language acquisition*. Hillsdale, NJ: Erlbaum.
So. L. K. H. & Dodd, B. (1995). The acquisition of phonology by Cantonese-speaking children. *Journal of Child Language* **22**, 473–95.
Stampe, D. (1973). *A dissertation on natural phonology*. Unpublished doctoral dissertation, University of Chicago.
Timm, L. A. (1977). A child's acquisition of Russian phonology. *Journal of Child Language* **4**, 329–39.
Topbas, S. (1997). Phonological acquisition of Turkish children: implications for phonological disorders. *European Journal of Disorders of Communication* **32**, 377–96.
Vihman, M. M. (1996). *Phonological development*. Oxford: Blackwell.
Wang, J. Z. (1989). *Mandarin syllable structure and its implications for phonemic analysis*. Unpublished doctoral thesis, University of Delaware.
Yasuda, A. (1970). Articulatory skills in three-year-old children. *Studia Phonologica* **5**, 52–71.
Yavaş, M. (1997). Feature enhancement and phonological acquisition. *Clinical Linguistics and Phonetics* **11**(2), 153–72.

APPENDIX 1

No.	English	Pinyin	IPA	No.	English	Pinyin	IPA
1	nose	bizi	pi2·tsi0	23	light	deng	tʌŋ1
2	ear	erduo	ɚ3·tuo0	24	umbrella	yusan	y3·san3
3	mouth	zui	tsuei3	25	sun	taiyang	t^{h}ae4·iɑŋ0
4	finger	shouzhi	ʂow3·t ʂi3	26	moon	yueliang	yə4·liɑŋ0
5	hair	toufa	t^{h}ow2·f ʌ0	27	star	xingxing	ɕiŋ1·ɕiŋ (ɹ)0
6	foot	jiao	tɕiɑo3	28	flower	hua	xuʌ(ɹ)1
7	shoe	xie	ɕiɛ2	29	bird	niao	niɑo(ɹ)3
8	skirt	qunzi	$tɕ^{h}$yn2·tsi0	30	panda	xiongmao	ɕyoŋ2·mɑo(ɹ)1
9	apple	pingguo	p^{h}iŋ2·kuo3	31	plane	feiji	fei1·tɕi1
10	watermelon	xigua	ɕi1·kuʌ0	32	car	xiaoqiche	ɕiɑo3·$tɕ^{h}$i4·t ʂ^{h}ɤ(ɹ)1
11	banana	xiangjiao	ɕiɑŋ1·tɕiɑo1	33	ball	qiu	$tɕ^{h}$iow2
12	meat	rou	ɻow4	34	piano	gangqin	kɑŋ1·$tɕ^{h}$in2
13	vegetable	cai	ts^{h}ae4	35	girl	nühai	ny3·xae(ɹ)2
14	bowl	wan	uan(ɹ)3	36	boy	nanhai	nan2·xxae (ɹ)2
15	chopsticks	kuaizi	k^{h}uae4·tsi0	37	red	hong	xoŋ2
16	knife	dao	tɑo1	38	heart	xin	ɕin1
17	table	zhuozi	tʂuo1·tsi0	39	thank you	xiexie	ɕiɛ4·ɕiɛ0
18	water	shui	ʂuei3	40	goodbye	zaijian	tsae4·tɕiɛn4
19	wash face	xilian	ɕi3·liɛn3	41	stick	gunzi	kuən (ɹ)4·tsi0
20	brush teeth	shuaya	ʂuʌ1·i ʌ2	42	book	shu	ʂu1
21	bed	chuang	tʂ^{h}uɑŋ2	43	clip	jiazi	tɕiA1·tsi0
22	gate	men	mən(ɹ)2	44	circle	yuanquan	yan2·$tɕ^{h}$yan (ɹ)1

* *Note*: Pinyin is Chinese Romanization system. The numbers used in IPA transcription are tone indicators, representing high level, high rising, falling-rising and high falling tone respectively. Weakly stressed syllable is marked by the number 0. Rhotacized feature is marked by parentheses.

https://doi.org/10.1017/S030500099900402X Published online by Cambridge University Press

APPENDIX 2

FREQUENCY DISTRIBUTION OF PHONEMES, TONES, TONE SANDHI, WEAK STRESS, AND RHOTACIZATION IN THE TEST

Syllable-initial consonants (total occurrence in the test = 57)

Phonemes	Frequency	Phonemes	Frequency
p	1	x	4
p^h	1	s	1
t	3	ts	2
t^h	2	ts^h	1
k	4	ɕ	9
k^h	1	tɕ	5
m	2	$tɕ^h$	5
n	3	ʂ	4
f	2	tʂ	2
l	2	$tʂ^h$	2
ɹ	1	—	—

Vowels (total occurrence in the test = 67)

Phonemes	Frequency	Phonemes	Frequency
i	12	ow	3
u	1	iA	5
y	3	iɛ	5
o	1	uA	5
ɤ	1	uo	3
A	4	yɛ	1
ə	6	iɑo	4
əɻ	1	iow	1
ae	5	uae	1
ei	1	uei	2
ɑo	2	—	—

Syllable-final consonants (total occurrence in the test = 21)

Phonemes	Frequency	Phonemes	Frequency
n	11	ŋ	10

Tones (total occurrence in the test = 58)

Tones	Frequency	Tones	Frequency
Tone 1	17	Tone 2	16
Tone 3	15	Tone 4	10

Tone sandhi (total occurrence in the test = 6)

Tone sandhi type	Frequency
Falling-rising tone becomes a rising tone if followed by another falling-rising tone	3
Falling-rising tone retains the falling part in its contour only without rising in the pitch if followed by high level, rising and falling tones	2
High falling tone becomes low falling tone before another high falling tone	1

Weak stress (total occurrence in the test = 13)

Weak stress type	Frequency
Noun suffixes are weakly stressed	6
Reduplicated second syllables are weakly stressed	2
Second lexemes in some compounds are weakly stressed	5

Rhotacization (total occurrence in the test = 7)

Rhotacization type	Frequency
Words which are always rhotacized	4
Words which are optionally rhotacized	3

https://doi.org/10.1017/S030500099900402X Published online by Cambridge University Press

APPENDIX 3

COMPARISON OF PUTONGHUA, CANTONESE, ENGLISH AMD XHOSA PHONOLOGY

	Putonghua	Cantonese	English	Xhosa
Tones	4	9	None	9
Vowels	9 monophthongs 9 diphthongs 4 triphthongs	11 monophthongs 11 diphthongs	12 monophthongs 9 diphthongs 5 triphthongs	7 monophthongs
Syllable-final consonants	2	8	21	None
Syllable-initial consonants	21	17	23	41
Clusters	None	2	49	Very Few
Syllable structure	[C]V[C]	[C] [G]V [C/G]	$[C_{0-3}]$ V $[C_{0-4}]$	CV

Adapted from So & Dodd, 1995.

https://doi.org/10.1017/S030500099900402X Published online by Cambridge University Press