Speech rhythm of monolingual and bilingual children at age 2;6: Cantonese and English*

PEGGY P. K. MOK
The Chinese University of Hong Kong

Previous studies have showed that at age 3;0, monolingual children acquiring rhythmically different languages display distinct rhythmic patterns while the speech rhythm patterns of the languages of bilingual children are more similar. It is unclear whether the same observations can be found for younger children, at 2;6. This study compared five Cantonese–English simultaneous bilingual children with five monolingual children in each language using both rhythmic metrics and qualitative data on syllable structure complexity and lexical stress. Results show that while the speech rhythms of monolingual children are different at 2;6, the rhythmic patterns of bilingual children are less distinct.

1. Introduction

This study investigates the speech rhythm development of Cantonese–English bilingual children at 2;6, and their age-matched monolingual peers. It is an extension of Mok (2011), who showed that at age 3;0, monolingual children acquiring English (stress-timed) and Cantonese (syllable-timed) display distinct rhythmic patterns while the differences between the two languages of the bilingual children are less distinct. Bilingual English has less durational variability than monolingual English. The phonological developmental trajectory of bilingual children is distinct from that of monolingual children, a distinction which is manifested in acquisition delay and is influenced by language dominance. In addition, Kehoe, Lleó and Rakow (2011) found similar patterns in the speech rhythm of monolingual and bilingual children at 3;0 acquiring Spanish and German: monolingual children are significantly different from each other, while bilingual children acquiring the same two languages were more similar. The main research question of the current study is whether such rhythmic patterns can also be observed at a younger age (2;6) for children acquiring Cantonese and English monolingually and bilingually.

It is important to first point out that speech rhythm is a general term covering many phonological aspects (Dauer, 1983). There is also a continuum between typical stress-timed (e.g. English and German) and typical syllable-timed (e.g. Cantonese and Spanish) languages. Languages like Welsh and Catalan can fall between these typical languages (e.g. Grabe & Low, 2002). In addition, speech rhythm should not be defined only in terms of duration, despite the popular use of durational metrics (e.g. Arvaniti, 2009; Cumming, 2009). The current study compared typical examples of stress-timed languages (English) and syllable-timed languages (Cantonese) using both durational metrics and qualitative data on other phonological aspects. It is hoped that the results can provide a more comprehensive picture of rhythmic development.

There are good reasons to study monolingual and bilingual rhythm at a younger age (2;6) although it is unlikely that one will find more rhythmic differences than older children at 3;0. First, there is no study on bilingual speech rhythm before age 3;0. Second, very few studies have investigated monolingual rhythm of European languages before age 3;0, and those few studies have mainly examined durational patterns. The current study is on Cantonese and English, two rhythmically and typologically different languages. Every new language pair that is investigated provides a more complete picture of the bilingual mind and its inner-workings. Moreover, the use of additional qualitative data on syllable structure complexity, stress and vowel reduction in this study can also widen our understanding of early rhythmic development.

We can expect to find interesting rhythmic patterns at a young age for several reasons. First of all, speech

Address for correspondence:
Department of Linguistics and Modern Languages, Leung Kau Kui Building, The Chinese University of Hong Kong, Shatin, Hong Kong
peggymok@cuhk.edu.hk

rhythm has an underlying, foundational importance in early language acquisition. Several studies have demonstrated that rhythmic differences of the mothers' languages allow newborn infants to distinguish languages (Bosch & Sebastián-Gallés, 2001; Mehler, Jusczyk, Lambertz, Halsted, Bertoncini & Amiel-Tison, 1988; Nazzi, Bertoncini & Mehler, 1998; Nazzi, Jusczyk & Johnson, 2000). This ability is attributed to the knowledge of native language prosody already gained in utero in the third trimester (Houston, 2011). If even newborns are sensitive to linguistic rhythmic patterns, it is reasonable to expect that rhythmic differences can be observed in young children, although the exact patterns in production still remain unclear, particularly given the large gulf between perceptual sensitivity to rhythm and children's ability to express rhythmic differences in production. Lleó & Kehoe (2002) also argued that it is important to assess bilingual phonological acquisition during the second year of life, especially in the area of prosody, because language-specific segments are generally acquired later. Prosodic properties provide a better window for phonological comparison.

Second, Allen and Hawkins (1980) suggested that children begin with a more syllable-timed rhythm regardless of their target languages, because two important elements of stress-timing – consonant clusters and vowel reduction – are difficult for children to master. It would be expected, therefore, to observe weaker stress-timing in younger children's English when compared to the syllable-timing of the younger children's Cantonese, which is more stable rhythmically and relatively easier to articulate. There will be less variability in younger children's English compared to that of older children.

Findings of three previous studies support our hypothesis of rhythmic differences at 2;6. Vihman, Nakai and DePaolis (2006) compared durational patterns of disyllabic babbling and identifiable disyllabic words or phrases produced by infants acquiring American English, French and Welsh monolingually at two developmental stages: four-word point (around 12 months) and 25-word point (around 17 months). The three languages differ in speech rhythm: English (stress-timed), French (syllable-timed) and Welsh (ambiguous between the two). They found that at 12 months, children's disyllabic word production did not yet match adults' rhythmic patterns. However, at around 17 months (1;5), the French and Welsh children generally conformed relatively closely to the adult patterns, while the English children were more variable and conformed less to the adult patterns. Vihman et al. (2006) suggested that the inherent rhythmic variability of stress-timing poses a phonetically more challenging model for children to acquire than syllable-timing does, with more phonotactically complex sequences and segments not commonly found in children's production repertoires. In addition, there was more

variability in the adult English input than that in French or Welsh. As a result, word production of the English children departed from the adult forms more than the French or Welsh children's did. They suggested that since children's early vocal production patterns are common across languages and are biomechanically constrained, the matching between these cross-language properties and rhythmically different ambient languages would pose different challenges for children to acquire.

More recently, Payne, Post, Astruc, Prieto and Vanrell (2012) showed that the rhythmic influence of the ambient language is evident from a very early age for monolingual children. They found that two-year-old monolingual children speaking English (stress-timed) and Spanish (syllable-timed) or Catalan (ambiguous rhythmically) already displayed significant differences in durational variability measured by the rhythmic metrics. English child speech is already more variable than Spanish and Catalan child speech at age two, although they also showed that durational variability of child speech in these three languages differed from that of adult speech in a way that cannot be simply categorized as being more syllable-timed. They argued that the rhythm of child speech exhibits both common properties across languages and language-specific properties at the same time, similar to the suggestions by Vihman et al. (2006).

While Vihman et al. (2006) and Payne et al. (2012) demonstrated that monolingual children acquiring rhythmically different languages had distinct patterns by 2;0, the rhythmic development of bilingual children younger than 3;0 is less clear as no study has specifically investigated this. However, Paradis (2001) compared speech patterns of monolingual and bilingual French–English children at around age 2;6 using a nonsense-word repetition task. She found that their omission/truncations of the target words displayed biases towards language-specific stress patterns, and that the truncation patterns of the bilingual children were not the same as those of monolinguals. Although she did not specifically compare speech rhythm using durational metrics, her findings are relevant to the present research question about speech rhythm development, since stress is an important feature of speech rhythm (Dauer, 1983). French, like Cantonese, is a syllable-timed language. If bilingual 30-month-old children are sensitive to language-specific stress patterns in word truncation, they may demonstrate sensitivity to different rhythms in utterances too, but the exact patterns await investigation.

Impressionistic observation suggests that monolingual English at 2;6 does sound more regular and is closer to syllable-timing than that of monolingual English children at 3;0, while monolingual and bilingual Cantonese at 2;6 sound more comparable rhythmically. It will be interesting to examine how bilingual English at 2;6 patterns with these counterparts. In addition to using rhythmic metrics,

Table 1. *Background information of the five bilingual children.*

Child	Native language of mother	Native language of father	Sex	Language dominance	Age range for data used
Timmy	Cantonese	British English	M	Cantonese	2;5.19–2;6.09
Sophie	Cantonese	British English	F	Cantonese	2;5.02–2;6.12
Alicia	Cantonese	British English	F	Cantonese	2;5.04–2;7.28
Llywelyn	Cantonese	British English	M	Cantonese	2;5.10–2;7.04
Charlotte	Cantonese	British English	F	English	2;5.19–2;7.23

Table 2. *Background information of the monolingual children.*

	Cantonese				British English		
Child	Age at recording	Sex	Source	Child	Age at recording	Sex	Source
KOM	2;6.01	F	HKU-70	Ella	2;5–2;7	F	Forrester
LYW	2;5.23	F	HKU-70	Madi	2;7.01	F	Twin
TUNG	2;5.14	F	HKU-70	Fletchy	2;7.01	M	Twin
LSM	2;6.18	F	Local	Hannah	2;5–2;6	F	England
LTH	2;5.10	M	Local	Charlie	2;6–2;7	M	England

qualitative data on syllable structure and stress pattern will also be used in this study to complement the patterns found using durational rhythmic metrics. The results of this study together with those of Mok (2011) can give a better understanding of the development of speech rhythm in both monolingual and bilingual children's production at two points in time (2;6 and 3;0) when prosody is developing rapidly.

2. Method

2.1 Participants

Data from five of the six simultaneous Cantonese–English bilingual children in Mok (2011) were used in this study; data from one child (Janet) could not be included because she was not recorded before the age of 2;10. Table 1 shows the background information of the five bilingual children. Yip and Matthews (2007) give detailed background of these children. They were the offspring of inter-cultural marriages who were exposed to Cantonese and English from birth and grew up in a 'one parent one language' environment. They were recorded longitudinally at weekly or bi-weekly intervals in two unstructured play situations, one for Cantonese and one for English, although some language mixing can be found in the recordings, especially the earlier ones. The recordings can be found in the Hong Kong Bilingual Child Language Corpus, which is available through the YipMatthews corpus in CHILDES (http://childes.psy.cmu.edu/media/Biling/YipMatthews/).

Speech materials from five monolingual English and five monolingual Cantonese children were used for comparison. Data of the monolingual children came from various sources (see Table 2). The first three Cantonese children in Table 2 are featured in the HKU-Cantonese-70 corpus available in CHILDES (Fletcher, Leung, Stokes & Weizman, 2000). The other two Cantonese children were recorded locally by the author. LTH was also included in Mok (2011). Materials for the first English child (Ella) came from the Forrester corpus in CHILDES (Forrester, 2002). Two English children were fraternal twins from an expatriate family in Hong Kong recruited through word of mouth and recorded by the author. These three English children were also included in Mok (2011). The other two English children were recorded in England. The contents of the monolingual recordings were similar to the bilingual ones: natural conversations recorded in unstructured sessions.

2.1 Materials

The same research protocol as in Mok (2011) was used in this study. Essential information will be given here (see Mok, 2011, for fuller detail). Some 20 or 30 utterances of between four and nine syllables within the same breath group with minimal pausing were used for each child (see Table 3). Shorter utterances were not used because they are unsuitable for calculating durational variability, while longer utterances were quite rare in the speech of the children. The choice of utterances depends on factors like the quality and length of the original recordings and how talkative the child was in the recordings. Also, utterances with unnatural intonation, excessive stress on a particular syllable or word, excessive final lengthening or too much

Table 3. *Number of utterances, averaged number of syllables per utterance, and averaged speech rate (syllables/second) for each child; standard deviations are shown in brackets.*

		CANTONESE				ENGLISH		
	Child	No. of utterances	Averaged syllable no.	Speech rate	Child	No. of utterances	Averaged syllable no.	Speech rate
Mono-lingual	KOM	26	6.4 (1.2)	4.0 (0.8)	Ella	31	5.3 (1.0)	3.0 (0.5)
	LYW	26	5.7 (1.5)	3.7 (0.7)	Madi	31	5.7 (1.1)	3.5 (0.6)
	TUNG	26	5.3 (1.3)	3.9 (0.6)	Fletchy	33	4.9 (1.1)	3.5 (0.6)
	LSM	23	4.7 (1.1)	4.0 (0.9)	Hannah	31	5.8 (1.2)	3.5 (0.6)
	LTH	33	5.6 (1.5)	3.7 (0.8)	Charlie	22	5.6 (1.5)	3.4 (0.7)
	Mean:	**26.8 (3.7)**	**5.6 (0.6)**	**3.8 (0.2)**		**29.6 (4.3)**	**5.5 (0.4)**	**3.4 (0.3)**
Bilingual	Timmy	27	5.7 (1.1)	3.5 (0.8)	Timmy	23	5.2 (1.4)	3.4 (0.5)
	Sophie	26	5.6 (0.6)	3.5 (0.7)	Sophie	25	5.9 (1.0)	3.0 (0.5)
	Alicia	22	5.9 (0.9)	3.3 (0.8)	Alicia	27	4.7 (0.8)	3.0 (0.6)
	Llywelyn	27	5.5 (0.5)	3.9 (0.7)	Llywelyn	30	5.7 (0.7)	3.5 (0.5)
	Charlotte	19	5.2 (0.9)	4.0 (0.8)	Charlotte	32	6 (1.4)	3.9 (0.7)
	Mean:	**24.2 (3.6)**	**5.6 (0.3)**	**3.7 (0.3)**		**27.4 (3.6)**	**5.5 (0.5)**	**3.4 (0.4)**

background noise were not chosen. Only interpretable natural utterances showing clear formant structure in the spectrograms were chosen.

Each utterance was segmented into intervals at three levels (consonantal, vocalic and syllabic intervals in milliseconds). Final syllables were not excluded because utterances with four syllables will be too short for calculating variability if the final syllables were excluded. In addition, Bunta and Ingram (2007) showed that the inclusion or exclusion of the final syllables did not affect their results using normalised rhythmic metrics. Segmentation criteria for consonantal and vocalic intervals were mainly acoustic. Segmentation of Cantonese syllables was straight forward because of its simple syllable structure (see more details in Section 3.2 below). For English syllables, the maximal onset principle was followed as long as it produced phonotactically permissible onsets. The contentious ambisyllabic intervocalic consonants, e.g. the *m* in *lemon*, were treated as the onsets of the second syllables.

The same nine rhythmic metrics in Mok (2011) were used in the present study: ΔC, VarcoC, rPVI-C, nPVI-C, VarcoV, nPVI-V, %V, VarcoS and nPVI-S, where 'C' stands for consonants, 'V' for vowels, 'S' for syllables, 'r' for raw measurements, and 'n' for normalised measurements. ΔC (standard deviation of consonantal duration), %V (percentage of vocalic duration in speech), and VarcoC, VarcoV and VarcoS (normalised versions of Δ measures) calculate global durational variability of the whole utterance (Dellwo, 2006; Ramus, Nespor & Mehler, 1999); PVI (Pairwise Variability Index) captures the pairwise differences in duration between successive

units, and thus is a measure of local variability (Grabe & Low, 2002). A higher value of the eight rhythmic metrics (except %V) shows more durational variability, which indicates characteristics of stress-timing. Due to frequent vowel reduction and the prevalence of consonant clusters, stress-timed languages will have a lower %V than syllable-timed languages.

Since Mok (2011) showed that syllable variability (VarcoS and nPVI-S) can best illustrate rhythmic differences in the speech of young children, we will focus on these two metrics in the Results section, while the results of the other metrics are also reported for comparison with other studies. Detailed discussion of the following two relevant methodological issues can be found in Mok (2011): (i) the rationale of the above nine rhythmic metrics; and (ii) why syllable variability is more robust and more appropriate than variability of either consonant and vowel intervals for young children.

In addition to rhythmic metrics, two types of qualitative data are used to illustrate the patterns obtained using the rhythmic metrics: occurrence of different syllable types and vowel durational ratio of disyllabic trochaic English words. Details about these qualitative data will be given in the Results section.

3. Results

3.1 Rhythmic metrics

The left panel of Figure 1 shows the comparisons between monolingual Cantonese and monolingual English children. As expected, monolingual English
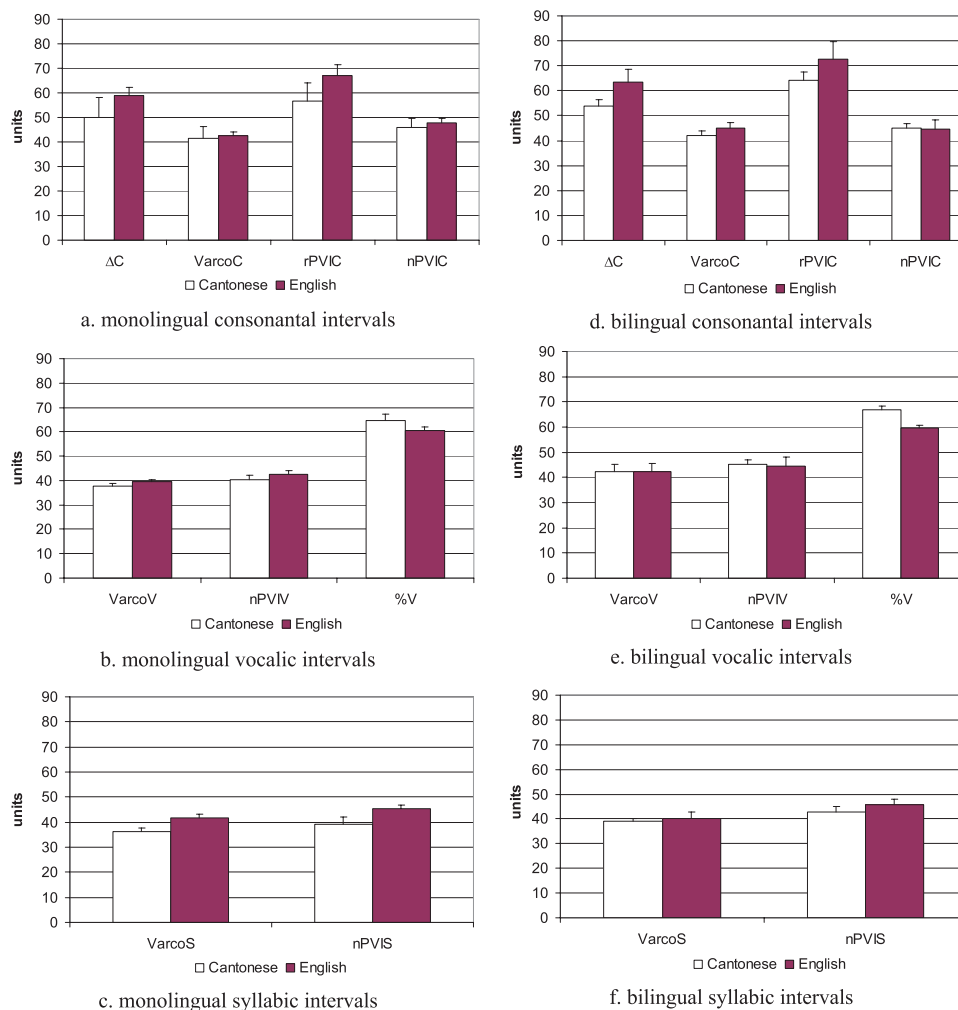
Figure 1. (Colour online) Rhythmic metrics of consonantal, vocalic and syllabic intervals of monolingual Cantonese and monolingual English children (left panel) and the two languages of the simultaneous bilingual children (right panel). The y-axis is relative to each rhythmic unit.

children had consistently higher values on the eight rhythmic metrics and a lower %V than monolingual Cantonese children. Detailed statistical results can be found in the Online Supplementary Materials. Most of the comparisons are not significant. Independent *t*-tests confirm that the difference is significant for VarcoS [$t(8) = -2.471$, $p = .039$] and approaching significance for nPVIS [$t(8) = -2.025$, $p = .077$]. The patterns of the syllabic intervals suggest that monolingual English children have a higher durational variability in their production than monolingual Cantonese children. These two groups of monolingual children display distinct rhythmic patterns at age 2;6. The comparisons between the two languages of the bilingual children are shown in the right panel of Figure 1. The patterns are very similar to those of the monolingual children. Nevertheless, paired sample *t*-tests show that the differences of the two syllabic metrics are not significant (VarcoS

[$t(4) = -0.638$, $p = .558$]; nPVIS [$t(4) = -0.941$, $p = .4$], although the difference in %V is significant [$t(4) = 8.802$, $p = .001$] (Cantonese having a higher %V than English, as expected). The results indicate that although bilingual children generally follow the monolingual patterns, the rhythm of their two languages are generally less distinct than that of their monolingual counterparts. Bilingual children have a developmental trajectory different from monolingual children. This concurs very well with the patterns of older children in Mok (2011).

In order to further examine the monolingual and bilingual patterns, we have put the VarcoS and nPVI-S values of the 2;6 children in a scatterplot (Figure 2A). We can see that the monolingual children are farther apart than the bilingual children. The bilingual values fall between the monolingual values. Independent *t*-tests confirm that there is no significant difference for the two metrics between monolingual Cantonese and
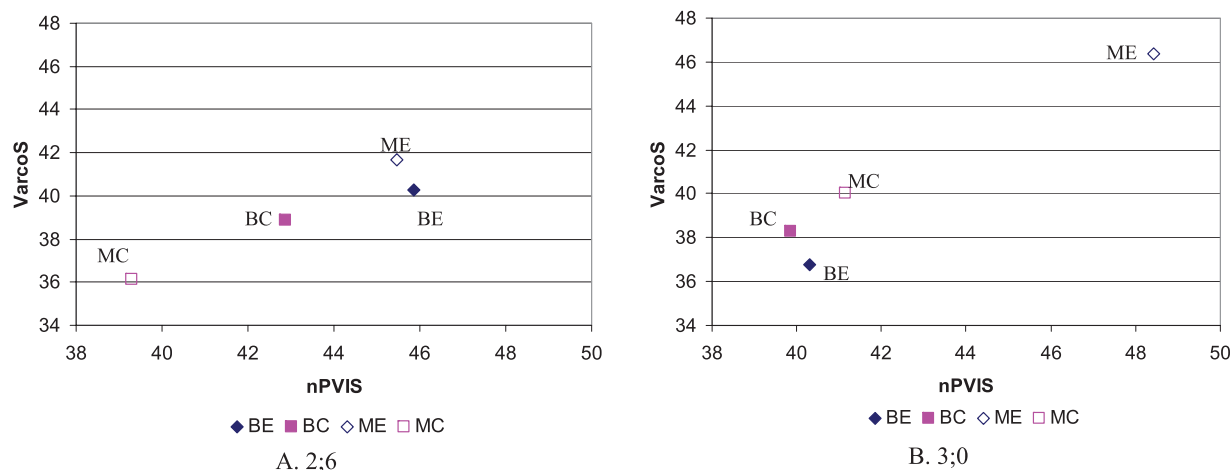
B = bilingual; M = monolingual; C = Cantonese; E = English

Figure 2. (Colour online) Scatterplots of syllabic intervals using VarcoS and nPVI-S at two ages: (A) 2;6, and (B) 3;0.

bilingual Cantonese, and between monolingual English and bilingual English, while the monolingual Cantonese and monolingual English children are significantly different, as reported above. These results again suggest that although the bilingual patterns are similar to the monolingual patterns, there is a larger rhythmic separation between the monolingual children than between the bilingual children at the same age.

The lack of significant difference between monolingual and bilingual English at 2;6 may give the impression that bilingual children's English is as stress-timed as that of the monolingual children. However, as pointed out by Allen and Hawkins (1980) and Vihman et al. (2006), stress-timing is challenging for young children to master. Young children acquiring stress-timed languages begin with a more syllable-timed rhythm. In order to better interpret the English patterns, the VarcoS and nPVI-S data of the three-year-old children in Mok (2011) were put in Figure 2B for comparison. It should be pointed out that while the data of the bilingual children at 2;6 and at 3;0 came from the same group of children (five at 2;6, six at 3;0), the data of the monolingual children are not all from the same children. Among the six monolingual children for each language at 3;0 (Figure 2B) and the five children for each language at 2;6 (Figure 2A), only one monolingual Cantonese and three monolingual English children were used at both ages. Also, the values were derived from uncontrolled natural utterances. Therefore, given these limitations, we should focus on their relative positions at the two ages separately rather than on the absolute values across ages, i.e., comparing the patterns within each figure but not comparing the same points across figures. Two obvious patterns can be observed if we compare the overall patterns in Figure 2A and 2B. First, monolingual English is further away in both metrics at 3;0 from the other children. Second, both languages of the

bilingual children cluster very closely with monolingual Cantonese at 3;0.

The simple comparison at the two ages confirms the difficulty of stress-timing for young children suggested by Allen and Hawkins (1980) and Vihman et al. (2006). Therefore, it is not that bilingual English at 2;6 is as STRESS-TIMED as monolingual English. Rather, monolingual English is closer to syllable-timing at 2;6. Syllable durational variability of monolingual English children has increased dramatically from 2;6 to 3;0 as compared with other children, which reflects the fact that stress-timing of monolingual English children developed considerably during this period. The additional Cantonese input between 2;6 and 3;0 has shaped the bilingual patterns to be closer to monolingual Cantonese.

3.2 *Qualitative data*

Syllable structure complexity, stress patterns and vowel reduction are important features of speech rhythm (Roach, 1982, Dauer, 1983). Two types of qualitative data were collected to illustrate the rhythmic differences between the two languages of monolingual and bilingual children in addition to the rhythmic metrics. The occurrence of different syllable types in Cantonese and English and the vowel durational ratio of disyllabic trochaic English words were examined for this purpose. All syllables produced by the monolingual and bilingual children in Cantonese and English were counted and checked auditorily. Each syllable was transcribed for its canonical syllable structure and how it was actually produced by the children in the recordings.

Morphological differences between Cantonese and English affect how syllable structure was counted in this study. The notion of 'word' is a complicated issue in Chinese (Packard, 2000). The phonological
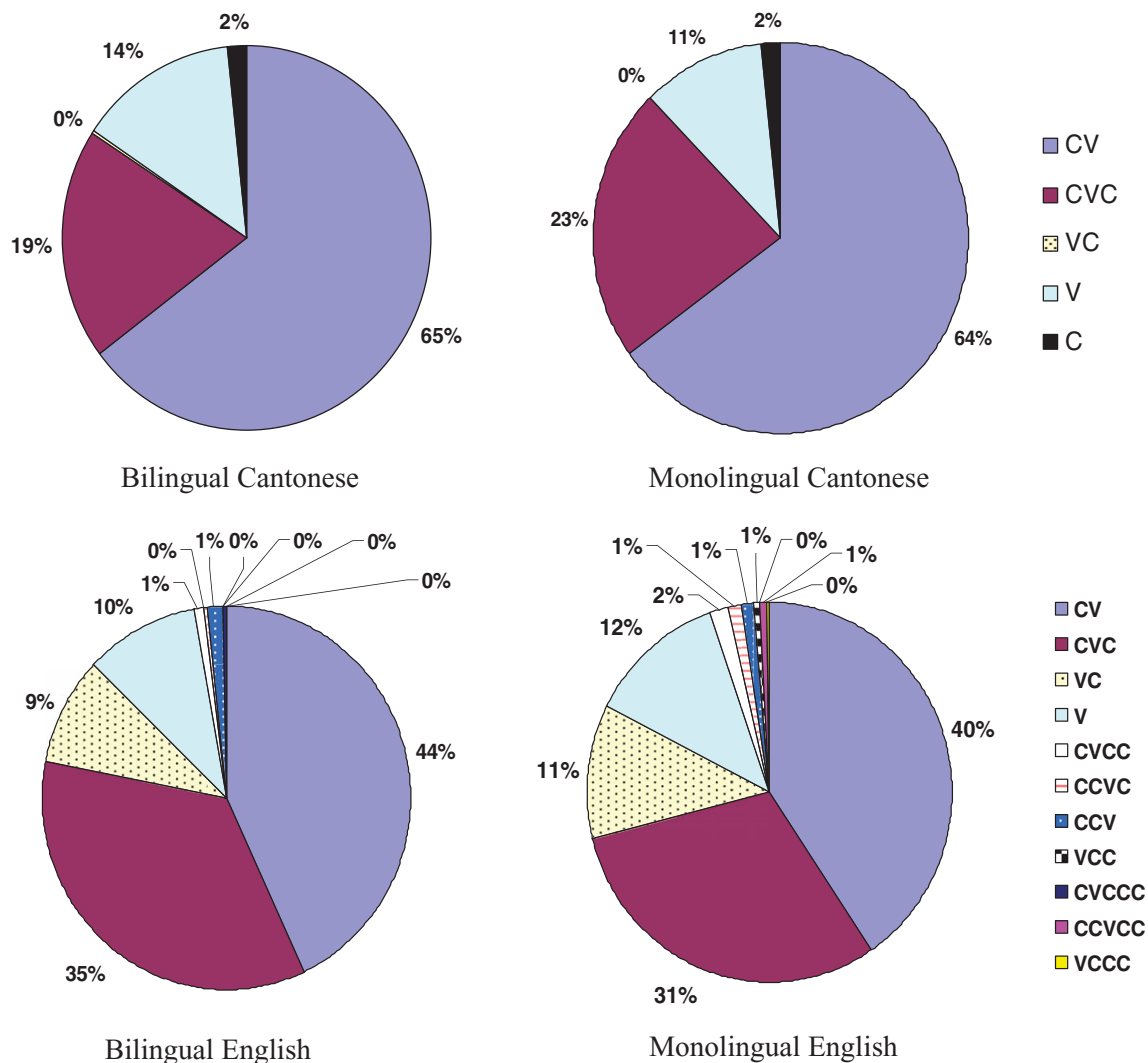
Figure 3. (Colour online) Occurrence (%) of different syllable structures in actual realisation produced by monolingual and bilingual children.

and morphological criteria for wordhood do not always coincide. The syllable is a very important unit in Chinese phonology, and a monosyllable can be a word or a morpheme. Morphologically disyllabic or multi-syllabic words in Chinese are always analysed as a sequence of monosyllables phonologically. Syllable structure in Cantonese is very simple: only /p t k m n ŋ/ can appear syllable-finally, and consonant clusters and resyllabification are prohibited. As a result, the segmentation of syllables in the Cantonese data was quite straightforward.

It is well understood that syllabication of intervocalic consonants is a tricky issue in English (e.g. Treiman & Danis, 1988; Treiman & Zukowski, 1990). Therefore, in order to have comparable data between Cantonese and English, all English words with more than one syllable were analyzed as a sequence of monosyllables, following

the practice in Chinese phonology. The maximal onset principle was followed as far as possible. The contentious ambisyllabic intervocalic consonants were treated as the onsets of the second syllables. For example, the word *melon* was considered to consist of two monosyllables: CV + CVC. This decision has probably led to an underestimation of the complexity of English syllables, but it allows comparable evaluation of syllable structure in the two languages.

Figure 3 shows the averaged percentage of different syllable types in actual realisation produced by the monolingual and bilingual children. Several observations can be made. First, Cantonese syllable structure is simpler than English structure for both monolingual and bilingual children. This alone is not surprising as it is a phonological difference between the two languages. What is more interesting is that among the simple and unmarked

Table 4. *Number of syllables involving consonant clusters in canonical form (and actual realisation) produced by the monolingual and bilingual English children.*

Syllable structure	Monolingual English					Bilingual English				
	Ella	Madi	Fletchy	Hannah	Charlie	Timmy	Sophie	Alicia	Llywelyn	Charlotte
CVCC	5 (0)	2 (2)	2 (0)	5 (2)	8 (6)	8 (0)	3 (2)	4 (0)	2 (1)	15 (3)
CCVC	1 (0)	0	0	1 (1)	3 (5)	0	4 (0)	1 (1)	0	0
CCV	1 (0)	1 (1)	2 (1)	2 (2)	3 (1)	4 (2)	4 (3)	1 (1)	2 (2)	0
VCC	1 (0)	4 (1)	2 (2)	1 (0)	1 (1)	0	6 (0)	5 (0)	1 (0)	2 (0)
CVCCC	0	0	0	0	0	0	1 (1)	0	0	1 (0)
CCVCC	1 (0)	0	0	0	5 (2)	0	2 (0)	0	0	0
VCCC	0	0	0	0	1 (1)	0	0	0	0	0

syllable types which occur in both languages (CV and CVC), Cantonese has a higher proportion of them than English, and the difference is more pronounced between monolingual children (Cantonese 87%, English 71%) than bilingual children (Cantonese 84%, English 79%). Second, only a few tokens of a marked syllable type, VC, are found in both monolingual and bilingual Cantonese (which averaged to 0%), while there is a much higher proportion in monolingual (11%) and bilingual (9%) English. It is worth mentioning again that the maximal onset principle was used in segmenting English syllables, so CVCVC structures were counted as CV + CVC. If we had used another strategy to segment the intervocalic consonants, e.g. CVC + VC, the above two differences between Cantonese and English syllable structure would be even more noticeable.

Third, the syllable types and frequency between monolingual and bilingual Cantonese are remarkably similar, whereas syllable structure in bilingual English is simpler than that of monolingual English. There are fewer complex syllable types and a higher proportion of unmarked syllable types (CV and CVC) in bilingual English (79%) than in monolingual English (71%). In order to further compare the syllable complexity between monolingual and bilingual English, the total numbers of syllables involving consonant clusters in both canonical form and actual pronunciation for each child are shown in Table 4. We can see that while the types of syllable structure attempted by both groups of children are similar, there is a higher incidence of structure simplification and reduction for bilingual children. Taken together, we can conclude that the phonological difference in syllable structure between Cantonese and English, which is an important feature of speech rhythm, is already clearly evident in the speech of monolingual and bilingual children at 2;6. The difference in syllable structure complexity is more pronounced in monolingual children than bilingual children. While bilingual children's Cantonese is comparable to their

monolingual counterparts, their English has simpler syllable structure than the monolingual children's. The findings on syllable structure complexity are in good agreement with the findings based on rhythmic metrics.

Stress and vowel reduction are also important features contributing to speech rhythm. The second type of qualitative data is the durational ratio of the two vowels in disyllabic trochaic words in English, which can reflect the mastery of both phonological features by the children. Disyllabic iambic words were not used because only very few iambic words were found in the recordings due to the strong trochaic tendency in English (Cutler & Carter, 1987). Despite this tendency in adult English, Vihman, DePaolis and Davis (1998) found that American English infants (around 1;5) also produced a fair amount of iambic disyllabic vocalizations, which could be attributed to the prevalence of iambic phrases (e.g. a ball) in the adult input. Such iambic phrases were also found in the children's recordings here, but they were excluded here so as to provide a simpler picture of stress acquisition. The first vowel in a disyllabic trochaic word should be longer than the second vowel. Since the speech materials were uncontrolled natural utterances, it is necessary to use a relative measure of vowel duration for valid comparison. Taking a ratio of the duration between the first and the second vowels in a disyllabic word can reasonably normalise the data for uncontrolled materials and individual differences.

Only disyllabic trochaic words in utterance medial position were used for two reasons. First, not every child produced usable disyllabic trochaic words in utterance initial position. Second, final lengthening would greatly distort the durational ratio of trochaic words in utterance final position. Table 5 shows the vowel durational ratios for each monolingual and bilingual child. We can see that the ratio is consistently higher than 1 for monolingual children, meaning that the trochaic words were produced with the first vowel longer than the second vowel. As for bilingual children, while some have a ratio similar

Table 5. *Durational ratios between the two vowels in disyllabic trochaic English words in utterance medial position.*

	Monolingual English			Bilingual English	
Child	No. of words	V1:V2	Child	No. of words	V1:V2
Ella	20	1.28	Timmy	3	1.14
Madi	14	1.32	Sophie	4	0.78
Fletchy	4	1.26	Alicia	5	1.31
Hannah	15	1.59	Llywelyn	15	0.97
Charlie	6	1.22	Charlotte	15	1.22
Average:		**1.34**			**1.08**

to monolingual children (e.g. Alicia 1.31), others have a ratio much lower than 1 (e.g. Sophie 0.78). Although the difference between the two groups of children is not significant [$t(8) = -2.171$, $p = .062$], there is a clear tendency for the monolingual children to have a stronger trochaic pattern than the bilingual children. Nevertheless, given the small numbers of disyllabic tokens for some children, caution should be exercised in the above interpretation. Clearly, further investigation is needed to corroborate this conclusion.

4. Discussion

Both the rhythmic metric data and qualitative data show that the speech rhythm of the monolingual children is different at 2;6: monolingual English and monolingual Cantonese possess features of stress-timing and syllable-timing, respectively. The two languages of the bilingual children at the same age generally follow the patterns of the monolingual children but are less distinct. Bilingual Cantonese is comparable with monolingual Cantonese, while bilingual English is less variable and simpler in structure than monolingual English. The monolingual and bilingual rhythmic patterns found for children at 3;0 in Mok (2011) are also present in the speech of children at 2;6.

The monolingual findings in this study concur well with Payne et al. (2012), who demonstrated that monolingual children at 2 acquiring rhythmically different languages have separate rhythms at an early age. Vihman et al. (2006) found that no clear language-specific rhythmic patterns were observable yet at four-word point (around 1;0) but differences were found at 25-word point (around 1;5). Based on all these results, we can assume that rhythmic separation between monolingual children begins during the short period between 1;0 and 1;5, and is different in terms of durational metrics by 2. Comparison between the current findings and those in Mok (2011)

suggests that the rhythmic patterns continue to diverge throughout the second year of life, with a noticeable increase in syllabic variability for monolingual English. The data in Payne et al. (2012) show that rhythmic divergence actually continues well into the following years.

The data in the current study, Kehoe et al. (2011), Lleó, Rakow and Kehoe (2007) and Mok (2011) suggest that the establishment of language-specific rhythms takes longer for bilingual than for monolingual children. Although bilingual children are already following the monolingual patterns at 2;6, the rhythm of the two languages as measured by the durational metrics is still not significantly different by age 3;0. Comparing monolingual and bilingual children acquiring Spanish and English, Bunta and Ingram (2007) found significant differences in durational metrics between the two languages of the bilingual children at around age of four. In addition, a case study of a balanced Cantonese–English bilingual child at ages three and four years also indicates that the two languages are more different at the age of four (White & Mok, 2011). Thus, we can infer that rhythmic divergence for bilingual children develops considerably in the fourth year of life. Further longitudinal studies at several time points between age three and age four are needed in order to track bilingual rhythmic development more closely. Differences depending on the target language pairs and individual children are likely to emerge.

Although bilingual speech rhythm is delayed compared to monolingual rhythm, the two languages are not simply following the monolingual patterns at a slower pace. Instead, they interact, resulting in a developmental trajectory distinct from that of monolinguals. Bilingual children are not simply two monolinguals in one (Grosjean, 1989). All of the studies on bilingual rhythm mentioned above show an asymmetry in that the stress-timed language is more affected by bilingual development than the syllable-timed language is. Not only is the durational variability reduced, syllable structure complexity is simpler and stress patterns are also weaker when compared to the monolingual stress-timed counterparts. The patterns of the syllable-timed language of the bilingual children are on a par with monolingual children. These results confirm the bias towards more equal timing and the difficulty of stress timing in early phonological acquisition (Allen & Hawkins, 1980; Vihman et al., 2006). In addition, they also show that bilingual phonological development is not delayed in an across-the-board manner. Some phonological features, e.g. those that are more challenging phonetically, are more vulnerable than others.

In addition to acquisition delay, language dominance can also account for the developmental patterns of bilingual speech rhythm (Genesee, Nicoladis & Paradis, 1995). Four of the five bilingual children in this study were

Cantonese-dominant, and they all lived in a Cantonese-speaking environment (Hong Kong) with Cantonese mothers (see Table 1). Their language dominance was measured objectively using mean length of utterance (MLU) and MLU differentials (the difference between MLU values for a child's two languages at a given time point), and ample evidence of the dominance of Cantonese syntactically can be found (see Yip & Matthews, 2007, for detail). Although their language dominance was measured by morphosyntactic complexity, the observed patterns in their speech rhythm match their language dominance very well. It is impossible to distinguish the effects of delay and language dominance separately based on the present data only. It is quite likely that both factors jointly contribute to their rhythmic patterns. Kehoe et al. (2011) showed that balanced German–Spanish bilingual children growing up in either Germany or Spain exhibit similar durational variability in their speech, and that both groups of bilingual children are less variable than their monolingual counterparts. Besides the findings on speech rhythm, Kehoe, Lleó and Rakow (2004) also found that the dominant language tends to influence the other language in early bilingual phonological acquisition of voice onset time (VOT). These results suggest that there is a close relationship between language dominance and the direction of cross-linguistic influence. It will be interesting to compare rhythmic development of bilingual children with dominance in either language to further investigate the effects of language dominance on early phonological acquisition.

Another possible interpretation of the rhythmic patterns is that the bilingual children may be settling on patterns that are in between their two languages. Similar cases can be found in the VOT values of adult bilingual speakers, who often do not exhibit monolingual-like VOT values in the target languages but produce values that are intermediate between those of their two languages (e.g. Flege, 1991; Laeufer, 1996). It is possible that during early rhythmic development, the bilingual children's production also exhibits a compromise between the two extremes, particularly given the fact that the two target languages (Cantonese and English) differ substantially in many phonological aspects. Faced with many phonological challenges at the same time, bilingual children may start out with intermediate values. Kehoe et al. (2011) also supported the idea that bilingual children may be employing phonetic compromise in their temporal realisations of consonants and vowels, resulting in similar rhythmic patterns in their German and Spanish. Nevertheless, bilingual children do not remain in the middle ground for long, unlike adult bilingual speakers. Bunta and Ingram (2007) showed that older bilingual children (4;6–5;2) had more rhythmic separation between Spanish and English than younger bilingual children (3;9–4;5), but they were still not the same as monolingual children. It is clear from their results that bilingual children continue to separate the two languages following the path of monolingual children. It will be interesting to see when bilingual children can acquire target-like rhythmic patterns similar to their monolingual counterparts, i.e., when bilingual children will, as it were, catch up with monolingual children. Whitworth (2002) suggested that bilingual speech rhythm is not completely acquired even by around age 11. Such a late acquisition age was suggested because she compared two languages in the same rhythmic group (German and English). Therefore, in addition to studies on young bilingual children, further studies on older bilingual children acquiring rhythmically different languages are also needed in order to fully understand their developmental trajectories.

Finally, although the speech rhythm of bilingual children younger than 2;6 is unlikely to be statistically significant in terms of durational metrics, this does not mean that it is futile to investigate the bilingual speech rhythm of even younger children, as demonstrated by Vihman et al. (2006). Qualitative data from the current study clearly demonstrate that phonological asymmetry exists even if the durational metrics reveal no significant differences. The findings caution against the use of durational metrics only in the study of speech rhythm, especially for young children. Subtle phonological differences are not captured by the durational metrics. Therefore, more detailed examination of individual phonological features is needed in future studies of bilingual rhythm of young children, rather than relying on the measurement of overall durational variability alone.

To conclude, the current study confirms previous findings on the monolingual speech rhythm of young children and provides new insights into bilingual speech rhythm development before age 3;0. It estimates the ages of rhythmic divergence for monolingual and bilingual children based on the current and several previous studies. More detailed investigation is needed to confirm these ages and to reveal the patterns of various phonological features contributing to such development.

References

Allen, G., & Hawkins, S. (1980). Phonological rhythm: definition and development. In G. Yeni-Komshian, J. Kavanagh & C. Ferguson (eds.), *Child phonology* (vol. 1): *Production*, pp. 227–256. New York: Academic Press.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica, 66,* 46–63.

Bosch, L., & Sebastián-Gallés, N. (2001). Early language differentiation in bilingual infants. In J. Cenoz & F. Genesee (eds.), *Trends in bilingual acquisition*, pp. 71–93. Amsterdam: John Benjamins.

Bunta, F., & Ingram, D. (2007). The acquisition of speech rhythm by bilingual Spanish- and English-speaking 4- and

5-year-old children. *Journal of Speech, Language and Hearing Research, 50 (4),* 999–1014.

Cumming, R. (2009). The interdependence of tonal and durational cues in the perception of rhythmic groups. *Phonetica, 67,* 219–242.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2,* 133–142.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11,* 51–62.

Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient for ΔC. In P. Karnowski & I. Szigeti (eds.), *Language and language-processing*, pp. 231–241. Frankfurt am Main: Peter Lang.

Flege, J. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America, 89,* 395–411.

Fletcher, P., Leung, S. C. S., Stokes, S. F., & Weizman, Z. O. (2000). *Cantonese preschool language development: A guide*. Hong Kong: Department of Speech and Hearing Sciences.

Forrester, M. (2002). Appropriating cultural conceptions of childhood: Participation in conversation. *Childhood, 9,* 255–278.

Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language, 22,* 611–631.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (eds.), *Laboratory phonology VII*, pp. 515–546. Berlin: Mouton de Gruyter.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36,* 3–15.

Houston, D. (2011). Infant speech perception. In R. Seewald & A. M. Tharpe (eds.), *Comprehensive handbook of pediatric audiology*, pp. 47–62. San Diego, CA: Plural Publishing.

Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German–Spanish children. *Bilingualism: Language and Cognition, 7,* 71–88.

Kehoe, M., Lleó, C., & Rakow, M. (2011). Speech rhythm in the pronunciation of German and Spanish monolingual and German–Spanish bilingual 3-year-olds. *Linguistische Berichte, 227,* 323–352.

Laeufer, C. (1996). The acquisition of a complex phonological contrast: Voice timing patterns of English initial stops by native French speakers. *Phonetica, 53,* 86–110.

Lleó, C., & Kehoe, M. (2002). On the interaction of phonological systems in child bilingual acquisition. *The International Journal of Bilingualism, 6,* 233–237.

Lleó, C., Rakow, M., & Kehoe, M. (2007). Acquiring rhythmically different languages in a bilingual context. *Proceedings of the 16th International Congress of Phonetic Sciences* (ICPhS 16), Saarbrueken, Germany, pp. 1545–1548.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29,* 143–178.

Mok, P. P. K. (2011). The acquisition of speech rhythm by three-year-old bilingual and monolingual children: Cantonese and English. *Bilingualism: Language and Cognition, 14,* 458–472.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 756–766.

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language, 43,* 1–19.

Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge: Cambridge University Press.

Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *The International Journal of Bilingualism, 5,* 19–38.

Payne, E., Post, B., Astruc, L., Prieto, P., & Vanrell, M. M. (2012). Measuring child rhythm. *Language and Speech, 55,* 203–229.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73,* 265–292.

Roach, P. (1982). On the distinction between stress-timed languages and syllable-timed languages. In D. Crystal (ed.), *Linguistic controversies: Essays in honour of F. R. Palmer*, pp. 73–79. London: Arnold.

Treiman, R., & Danis, C. (1988). Syllabification of intervocalic consonants. *Journal of Memory and Language, 27,* 87–104.

Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language, 29,* 66–85.

Vihman, M., DePaolis, R., & Davis, L. (1998). Is there a 'trochaic bias' in early word learning? Evidence from infant production in English and French. *Child Development, 1998,* 935–949.

Vihman, M., Nakai, S., & DePaolis, R. (2006). Getting the rhythm right: A cross-linguistic study of segmental duration in babbling and first words. In L. Goldstein, D. H. Whalen & C. T. Best (eds.), *Laboratory phonology 8*, pp. 343–368. New York: Mouton de Gruyter.

White, D., & Mok, P. P. K. (2011). A case study of speech rhythm acquisition in a Cantonese–English bilingual child. Presented at the International Congress for the Study of Child Language (IASCL), Montreal.

Whitworth, N. (2002). Speech rhythm production in three German–English bilingual families. *Leeds Working Papers in Linguistics and Phonetics, 9,* 175–205.

Yip, V., & Matthews, S. (2007). *The bilingual child: Early development and language contact*. Cambridge: Cambridge University Press.