The Setswana speech rhythm of 6-7 years old Setswana-English bilingual children

Article

Updated Version

Title: The Setswana speech rhythm of 6-7 years old Setswana-English bilingual children

Authors:

Boikanyego Sebina, University of Botswana

Jane Setter, University of Reading, UK

Michael Daller, University of Reading, UK

Abstract:

Aims and objectives

This study investigates the acquisition of Setswana speech rhythm, considered to be typically syllable-timed, by early sequential Setswana-English bilingual children aged 6-7 years old growing up in Botswana, a country with a diglossic setting, where English is the dominant high-status language in educational and public contexts. For this group of children, taught full-time in English from the age of three years, the L2 becomes their dominant language through exposure to English-medium education. The aim was to ascertain if the prosodic patterns of Setswana spoken by the bilingual children are similar to those of the monolingual children or if English, considered to be stress-timed, has an effect on these prosodic features.

Data and analysis

The speech rhythm patterns of 10 Setswana-English bilingual children are compared with those of 10 age-matched Setswana monolingual children educated in public schools for whom English is a learner language. The study primarily examines spontaneous speech from the telling of a wordless picture storybook, and utilises rhythm metrics nPVI-V and Varco V to examine the speech rhythm of the children.

Findings

The results showed that the prosodic pattern of Setswana in the bilingual group diverged from that of the non-bilingual group.

Originality

This is the first such study on speech rhythm in bilingual children in Setswana.

Significance

The research provides evidence in this population of effects from English bilingualism on L1 Setswana speech prosody, and challenges the assumption that speech rhythm prosody is established early in life, especially when the language is a less marked, syllable-timed language like Setswana.

Keywords: speech rhythm, bilingual, children, Setswana, English.

Introduction

Speech rhythm

The discussion of the concept of speech rhythm has been on going as far back as the 1960's. The study of speech rhythm has historically worked on the premise that languages can be classified into distinct rhythmic classes such as stress-, syllable- and mora timed (Abercrombie, 1967; Pike, 1945). Rhythm in speech relies on the notion of isochronous recurrence of some units in speech timing (Arvaniti, 2012; Grabe & Low, 2002). Extensive instrumental research failed to provide acoustic evidence for isochrony of the rhythmic units of feet in stress-timed languages and syllable duration in syllable-timed languages, however; Roach (1982) found that inter-stress intervals are not more equal in stress- than in syllable-timed languages, and Dauer's (1983) comparison of inter-stress intervals across languages did not indicate differences. Dauer (1983) concluded that languages exist on a rhythmic continuum from least stress-timed to most stress-timed. Roach (1982) and Dauer (1983) concluded that no language is explicitly stress-timed or syllable-timed; both types of timing are inherent in all languages. What varies is the degree of timing a language exhibits.

The lack of physical isochrony obliged researchers like Roach (1982) and Dauer (1983) to turn to a phonological account to explain why a given language may sound more stress-timed than another. Both suggested that the difference between rhythm classes in languages is due to the distinction in their syllable structures, vowel reduction and stress. Languages considered syllable-timed, like Setswana, have simple syllable structure; the majority of their syllables are open structure syllables (V, CV, CVV) and they do not generally exhibit consonant clusters. So-called stressed-timed languages, such as English, display a closed syllable structure (VC, CVC, CVCC) and allow consonant clusters. Vowels in stressed-timed languages are reduced, and at times are even absent in unstressed syllables. Vowel reduction is not a feature of syllable-timed languages, however. Regarding stress, syllable-timed languages tend not to have lexical stress while stress-timed languages do. Dauer's (1983) and Roach's (1982) findings resulted in a wealth of research on the dichotomy of languages in terms of rhythmic classes. Rhythm metrics resulted from these studies.

Rhythm metrics

Rhythm metrics are mathematical formulas used to classify languages into the rhythmic classes of stress-timing, syllable-timing and mora-timing (Arvaniti, 2012). The findings of Ramus, Nespor, Mehler (1999) inspired Grabe and Low (2002) to develop the Pairwise Variability Index (PVI), which indicates durational variability in successive vowel or consonant intervals. It is divided into two scores: raw (rPVI), for intervocalic intervals; and normalised (nPVI /nPVI-V), for vocalic intervals, the normalisation being applied for speech rate, because vowel durations are strongly affected by it. The PVI relies on the notion that the scores would be low when successive vowels or consonants in speech are similar. Therefore, because syllable-timed languages have open syllable structure that results in less varied syllables, they have lower PVI values compared to stress-timed languages, which exhibit variation in syllable structure and length. The most commonly applied formula, the normalised PVI (nPVI), is given in example (1), where *m* is the number of items in an utterance and *d* is the duration of the *k*th item:

(1)
$$\mathrm{nPVI} = 100\left[\sum_{k=1}^{m-1}\left|\frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2}\right| / (m-1)\right]$$

Variants of Ramus et al.'s (1999) metrics and the nPVI were developed to capture the rhythmic differences across languages. Of these variants, is the Varco V (White & Mattys, 2007a), which is a version of Ramus et al. (1999)'s ΔV normalised for speech rate. Varco V is the standard deviation of vocalic interval duration divided by mean vocalic interval duration multiplied by 100. The formula for Varco V is as follows (2):

(2)

$$\mathrm{VarcoV} = \frac{\Delta V}{\Delta} \times 100$$

Since the standard deviations in ΔV and the nPVI each measure different aspects of durational variability – that is, global and local variability respectively – it is important to apply both for an all-inclusive investigation, because there is a possibility of an utterance scoring high in global variability but low in local variability (Mok, 2011). Global durational variability (Varco V) measures the whole utterance while local durational variability (nPVI) measures differences between successive units (Mok, 2011). Similar to the nPVI, a higher value of Varco V is indicative of stressed-timed language due to more durational variability, while a low value would be expected for a syllable-timed language (Mok, 2011; White & Mattys, 2007a).

The reliability of rhythm metrics

Despite the development of Ramus et al.'s (1999) and Grabe and Low's (2000) metrics in an attempt to curb the discrepancies inherent in them, the problem of rhythm metrics not yielding similar results across different studies persisted (Arvaniti, 2009; Gut, 2012; White & Mattys, 2007b). Some studies suggested that the unreliability of rhythmic metrics could largely be attributed to material selection, speaking style, the segmentation techniques of the acoustic signal employed by different studies, as well as the elicitation of spoken data and the syllable composition of the material (Gut, 2012; Arvaniti, 2012). Nevertheless, vowel-based metrics like the nPVI, Varco V and %V, were found to be robust in differentiating languages into rhythmic classes compared to intervocalic metrics (Bunta & Ingram, 2007; Knight, 2011; Mok, 2011); they remain the most popular means of classifying languages into rhythm classes. The present study has adopted the nPVI and Varco V metrics to test the assumption that rhythm metrics are valid and reliable in discriminating languages and varieties as well as quantifying speech rhythm. The nPVI's and Varco V's measures of local and global durational variability respectively and the robustness of vocalic metrics in distinguishing languages into rhythmic classes has motivated their adoption in this study.

Acquisition of speech rhythm

Research on the acquisition of speech rhythm by children can be credited to the pioneering works of Allen and Hawkins (1980), who established that a child's first acquired speech rhythm is more syllable-timed regardless of the rhythm of their target language. Allen and Hawkins (1980) asserted that vowel reduction and consonantal clusters, which are the main components of stress-timed languages, are difficult for children to master. This hypothesis was tested by Grabe, Post, Watson (1999), who compared the nPVI scores of four-year-old monolingual English (stress-timed) and French (syllable-timed) speaking children with that of their mothers. The results supported Allen and Hawkins' hypothesis, as the rhythmic patterns of French children were similar to their mothers', while that of English children differed from that of their mothers, tending towards syllable timing.

Having established the rhythmic patterns of monolingual children acquiring stress-timed and syllable-timed languages, scholars set out to find out the rhythmic development of bilingual children acquiring languages that are both considered stress-timed, or both syllable-timed, or where one language is stress-timed and the other syllable-timed. Whitworth (2002) investigated speech rhythm patterns in six German-English bilingual children from the ages of 5-13 years using rPVI and nPVI. Since German and English are both considered stressed-timed, it is unsurprising that she did not find significant differences between the patterns in the children's two languages. Bunta and Ingram (2007) took a slightly different approach by investigating the development of rhythm in bilingual children from the ages of 3;9 - 5;2 years who were acquiring different rhythmic languages, Spanish and English, where Spanish is considered syllable-timed and English stress-timed. They found a significant difference between the rPVI and nPVI of the bilingual languages. The results also indicated significant results in the nPVIs of bilingual languages compared with monolinguals. The nPVI of bilingual English children indicated a lower variability compared to that of monolingual English children, whereas the difference between bilingual and monolingual Spanish children was not statistically significant. The findings are in support of Allen and Hawkins (1980) and Grabe et al. (1999), that young children's rhythm acquisition tends towards syllable-timing. The implication of Bunta and Ingram (2007) is that bilingual children as young as 4 years of age are able to keep the rhythmic classes of their two languages separate. However, it is possible that the Spanish bilingual and monolingual children were similar because syllable timing is easy to acquire, and so the debate continued on rhythmic development of bilingual children acquiring languages with different rhythm.

In a similar study to Bunta and Ingram (2007), Lleó, Rakow and Kehoe (2007) compared the rPVI and nPVI of three-year-old German-Spanish bilingual children to that of the same number and age of monolingual child speakers of these languages. The findings showed that while the German monolinguals' rhythmic patterns were different from that of the Spanish monolinguals, indicating that the two languages belong to different rhythm classes, the bilingual children displayed similar patterns in both of the languages. Contrary to Bunta and Ingram (2007), the results of the bilingual group in Lleó et al. (2007) showed that bilingual children merge the rhythm patterns of their languages, i.e., they do not keep them separate, as was the case with Bunta and Ingram (2007).

These contradictory results might be due to the ages of the participants. Studies have shown that rhythm is acquired quite early in childhood (Nazzi, Bertoncini, Mehler, 1998; Nazzi, Jusczyk, Johnson, 2000); therefore it is not surprising that children in Bunta and Ingram's study, who were older than those in Lleó et al. (2007), were able to keep the rhythm classes of the two languages separate, as they had been exposed to the two languages for a longer time. Interestingly, a study by Kehoe, Lleó, Rakow (2011) who used similar participants as in Lleó et al. (2007), showed that the bilingual children displayed similar rhythmic patterns in both languages indicating a bi-directional influence from both languages. Mok (2011) also established similar results with three-year-old simultaneous Cantonese-English bilinguals.

Likewise, Ordin & Polyanskaya (2014) investigated the development of speech rhythm in L1 and L2 by children and adults. A comparison of rhythmic patterns at a range of ages in L1 acquisition and at diverse proficiency levels in L2 showed that speech rhythm develops from syllable timing to stress timing in both groups. Further support for the conclusion that speech rhythm develops from a low vocalic variability is provided by Ordin and Polyanskaya (2015), who examined the development of speech rhythm in L2 acquisition by comparing vocalic variability in the English of German and French adult learners. The results of both groups of learners showed that as their proficiency increased, the vocalic variability also increased. Ordin and Polyanskaya (2015) concluded that acquisition of English speech rhythm by bilinguals develops from syllable timing to stress timing regardless of whether the language being acquired has similar rhythm timing with the native language of the learners. English and German are both considered stress-timed while French is considered syllable-timed.

The above studies on the rhythmic patterning of bilinguals give reason to investigate the rhythm patterns of older child participants (6-7 years) growing up in a diglossic environment, whose L1 is considered syllable-timed and whose L2 (English), which becomes dominant, is considered stress-timed. It will be valuable to our understanding of the processes involved to find out if the children keep their two languages separate, as in Bunta and Ingram's (2007) study, merge the two languages, as in Kehoe et al. (2011) and Mok (2011), or if the rhythm pattern tends towards the language environment, as in Kehoe and Lleó (2005). In addition, as far as we are aware, there has been no research in this area on Bantu languages such as Setswana.

The present study

The present study investigates the Setswana speech rhythm timing in the production of 6-7-year-old Setswana-English bilingual children who are exposed to high English input, including English-medium learning and a syllabus for English drawn from Cambridge Assessment International Education which uses a British English audio model and materials. The speech rhythm of this group of children is compared to age-matched Setswana monolingual children for whom English is considered a learner language. To the best knowledge of the researchers, the rhythm timing of Setswana has never been investigated in children. Like most Bantu languages, Setswana, the national language of Botswana, is considered syllable-timed (Coetzee & Wissing, 2007). The aim of this study is to find out if the speech rhythm of the bilinguals is similar to that of monolingual children.

An investigation of the speech rhythm of a language necessitates a discussion of the vowel system of that language. Setswana has seven vowels, short vowels (e.g., ɪ, i, ɛ, a, ɔ, ʊ, u) that consist of a single mora, while penultimate vowels are long vowels of the same quality, with two moras (e.g. ɪɪ, ii, ɛɛ, aa, ɔɔ, ʊʊ, uu) (Cole, 1955). Setswana does not have diphthongs. The penultimate syllable vowel length is an essential feature of the Setswana vowel system as it plays an important role in the variability of vowel duration (Cole, 1955; Sebina, Setter & Wright, 2019).

The study answers the following research questions:

1. What is the pattern of rhythm timing of Setswana in the speech of Setswana-English bilingual children aged 6-7 years in comparison with monolingual peers?
2. In the bilingual Setswana-English population, to what extent will the children in Standard (grade) 1, aged 6 years, have a different pattern of speech rhythm timing in Setswana in comparison with the children in Standard 2, aged 7 years, who will have had increased exposure to English?

The study makes the assumption that Setswana-English bilingual children will have a more stress-timed rhythm in their Setswana speech than their monolingual peers as a result of the influence of the British English model and materials used in private schools in Botswana; British English is considered to be stress-timed. Since there is no empirical study on the rhythm timing of the English spoken in Botswana – Coetzee and Wissing (2007) did not undertaken an empirical study on Tswana Englishes in general or Botswana English in particular – the hypothesis is based on the Speech Learning Model (Flege, 2002) which postulates that the L1 and L2 of bilinguals mutually influence each other, and the results of Kehoe et al. (2011) and Mok (2011), cited above, which demonstrate this to be the case. It also assumes that bilingual children in Standard 2, who have had longer exposure to British English, will display greater stress-timing in Setswana than their younger peers in Standard 1. It should be noted, however, that the English spoken by Batswana teachers (citizens of Botswana are referred to as "Batswana") in both contexts is likely to be syllable-timed to a lesser or greater extent.

Methods

Participants

The study employed twenty 6-7-year-old children with no speech and language impairment, ten of whom were Setswana monolinguals. The monolinguals are considered as such because they were in the first and second year of public primary school, where English is not the main medium of instruction. Therefore, they were exposed to minimal English. The remaining 10 are sequential Setswana-English bilinguals who have been exposed to high English input through private English medium schools from the age of three. All the children were born in Botswana and had never lived outside the country (see Tables 1 and 2). The children were recruited from schools through consent letters that were given to the children to give to their parents. Therefore, the children's participation in the study depended on the parents. The children were also given The Language and Social Background Questionnaire (LSBQ), adapted from (Bialystok, 2011) to be completed by the parents. The questionnaire elicited the language use pattern of the participants as this determined their inclusion in the study. In addition, the Raven's Coloured Progressive Matrices (RCPM) (Raven, 1982) was administered to ensure that the two groups did not cognitively differ significantly ($t(18)=2.003$, $p=.060$).

The 6-7-year age group was chosen based on the results of previous studies, which reported that by the age of 3 monolingual children display distinct rhythmic patterns for their language (Bunta & Ingram, 2007; Mok, 2011). Therefore, at the age of 6-7-years, the Setswana speech rhythm of these children should be in place.

Table 1 Background of Setswana monolingual children.

Participants	Age (years; months)	Gender	Standard (grade)	Country of birth	Age (years) started school
1	6.05	Male	1	Botswana	3
2	6.06	Female	1	Botswana	3
3	6.07	Female	1	Botswana	3
4	6.09	Female	1	Botswana	3
5	6.09	Male	1	Botswana	4
6	6.11	Female	2	Botswana	3
7	7.02	Female	1	Botswana	3
8	7.09	Male	2	Botswana	3
9	7.10	Female	2	Botswana	2
10	7.10	Female	2	Botswana	2

Table 2 Background of Setswana-English bilingual children

Participants	Age (years; months)	Gender	Standard (grade)	Country of birth	Age (years) started school
1	6.04	Female	1	Botswana	3
2	6.04	Female	1	Botswana	3
3	7.02	Female	1	Botswana	3
4	7.03	Male	1	Botswana	3
5	7.05	Female	2	Botswana	3
6	7.07	Male	2	Botswana	3
7	7.07	Male	2	Botswana	3
8	7.08	Female	2	Botswana	3
9	7.10	Female	1	Botswana	2
10	7.11	Male	2	Botswana	3

Material

This study mainly uses spontaneous speech based on the wordless picture storybook *Frog where are you?* (Mayer, 1969) (hereafter *Frog Story)* in order to collect rich sets of comparable data produced in a naturalistic context. The *Frog Story* has been extensively used in linguistic research (see Bayram & Wright, 2016; Reilly, Losh, Bellugi & Wulfeck (2004); Setter, Stojanovik, Van Ewijk, & Moreland, 2007) to determine the linguistic performance of participants. Since the book does not have words, it allows for a plentiful production of spontaneous language.

Procedure

Data collection

The procedure used in the present study to administer the storybook followed that of Reilly et al. (2004) (see Reilly et al., 2004). The children were audio recorded while they told the story. The children were acquainted with the recording device during the informal chat to ensure a relaxed atmosphere. The task took an average of 10 minutes. The task took place at schools in one of the quiet classrooms.

Recording instruments

A Roland Edirol R-09HR recording device collecting data in wave format sampled at a rate of 44.1kHz, 16-bit stereo was used to collect data of the best possible quality and facilitate subsequent analysis. The recordings were saved on a memory card, which was installed in the recorder, and later transferred to a computer (MacBook Air) for analysis.

Acoustic data analysis

Praat (Boersma & Weenink, 2007) was employed in the acoustic analysis of audio data. The researcher listened to the recordings on Praat and selected 60 seconds of speech with minimal pauses that was audible and comprehensible for rhythm calculation. It is worth noting that other studies on speech rhythm in bilingual children made use of between 20 to 30 intonational phrases / utterances of not less than five syllables for the analysis (see Mok, 2011; Mok, 2013;

Kehoe et al., 2011) while others used sentences that the children read out (see Bantu & Ingram, 2007). An intonational phrase is defined as “a sense group separated by a pause and forming a prosodic whole” (Kehoe et al., 2011, p. 334). The positioning of intonation phrase boundary is a controversial issue where even native speakers of a language can differ about its placement (Grabe & Low, 2002). Since it seems there is nothing in the literature about the Setswana intonation phrase boundary, it was decided that it was safe to use continuous monologue of around 60 seconds long, as this has been successfully applied in other studies with child participants (see Setter, Stojanovik, Ewijk, Moreland, 2007) as well as with adults (see Arvaniti, 2012). Grabe et al. (1999) suggested that utterances shorter than five syllables are not suitable for calculating the PVI. As a result of this, a number of utterances were selected from each child, none of which was less than five syllables long, and used for analysis in the present study.

The sound files were labelled segmentally by the main researcher through simultaneously listening to the recording and inspecting the waveform and spectrogram generated from Praat. The labelling was divided according to three tiers inserted on the Praat display, where Tier 1 is the annotation or transcription of an utterance. The utterances were transcribed orthographically into Setswana and glossed in English by the main researcher. Tier 2 was used to label the vowels, and Tier 3 was used for the different syllables of the word. The vowel durations were extracted from Praat and entered into a Microsoft Office Excel spreadsheet in preparation for the rhythm calculations. The measurement and segmentation criteria for vocalic intervals (vowels) followed that of Grabe and Low (2002), which are based on those of Peterson and Lehiste (1960).

To ensure reliability of the measurements, an individual trained in acoustic phonetics independently labelled and measured vowel duration from randomly selected 20% of the data, following a discussion about the vowel measurement methodology adopted in this research. A Spearman's Rho correlation test indicated a significant strong positive relationship between the original measurements and those of the independent rater (Spearman's Rho: $rs = .879$; $p < .001$). Any issues arising from the labelling and measurement procedures during the discussion of the methodology adopted in the present study were discussed fully before the independent labelling and measurement took place.

Results

The first research question is to determine the pattern of rhythm timing of Setswana in the speech of Setswana-English bilingual children aged 6-7 years in comparison with monolingual peers. The results of the means as indicated in Figure 1 show that the bilingual group has a lower nPVI-V and Varco V compared to the monolingual group. The independent samples *t*-test showed a statistically significant effect, nPVI-V [$t(18) = 6.284$, $p = .001$]; Varco V [$t(18) = 3.179$, $p = .005$] with an effect size of $d = 2.81$ for the nPVI-V and $d = 1.42$ for Varco V. Therefore, the bilingual group was associated with a statistically significantly smaller nPVI-V and Varco V value than the monolingual group. This was not the expected result based on previous research regarding acquisition of syllable-timed languages.

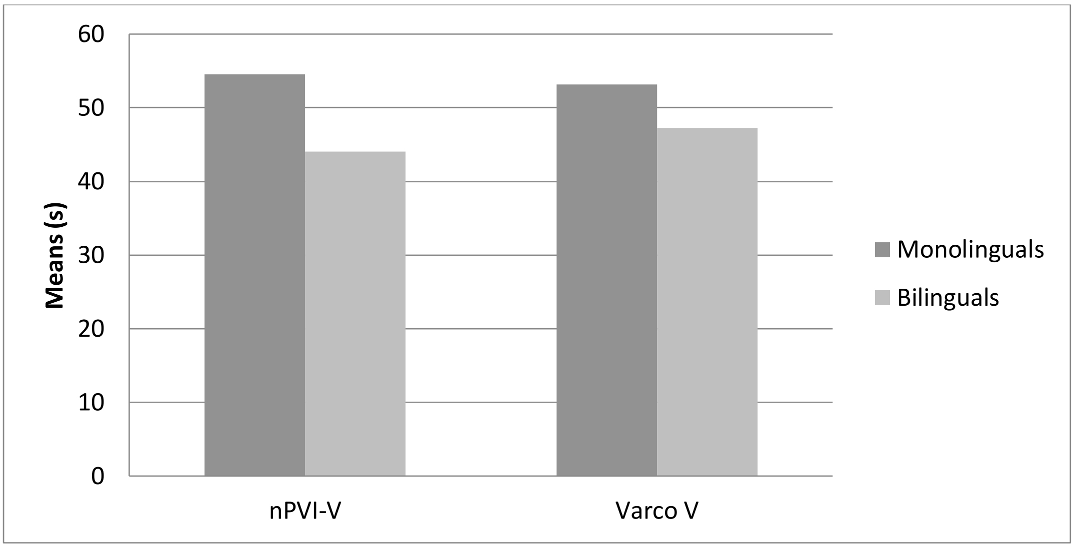

Figure 1. nPVI-V and Varco V means for monolinguals and bilinguals.

A Pearson correlation test was run to determine if there was a statistically significant relationship between the nPVI-V and Varco V variables. The results (Figure 2) are Pearson *r*: $r = .783$; $p < .001$, indicating a significant strong positive relationship between the nPVI-V and Varco V. The relationship is graphically represented by the scatterplot in Figure 2[1].

[1] At the request of our reviewers, we also carried out multiple regression analyses on the data. These did not result in any significant predictor variables or add to the results presented here.

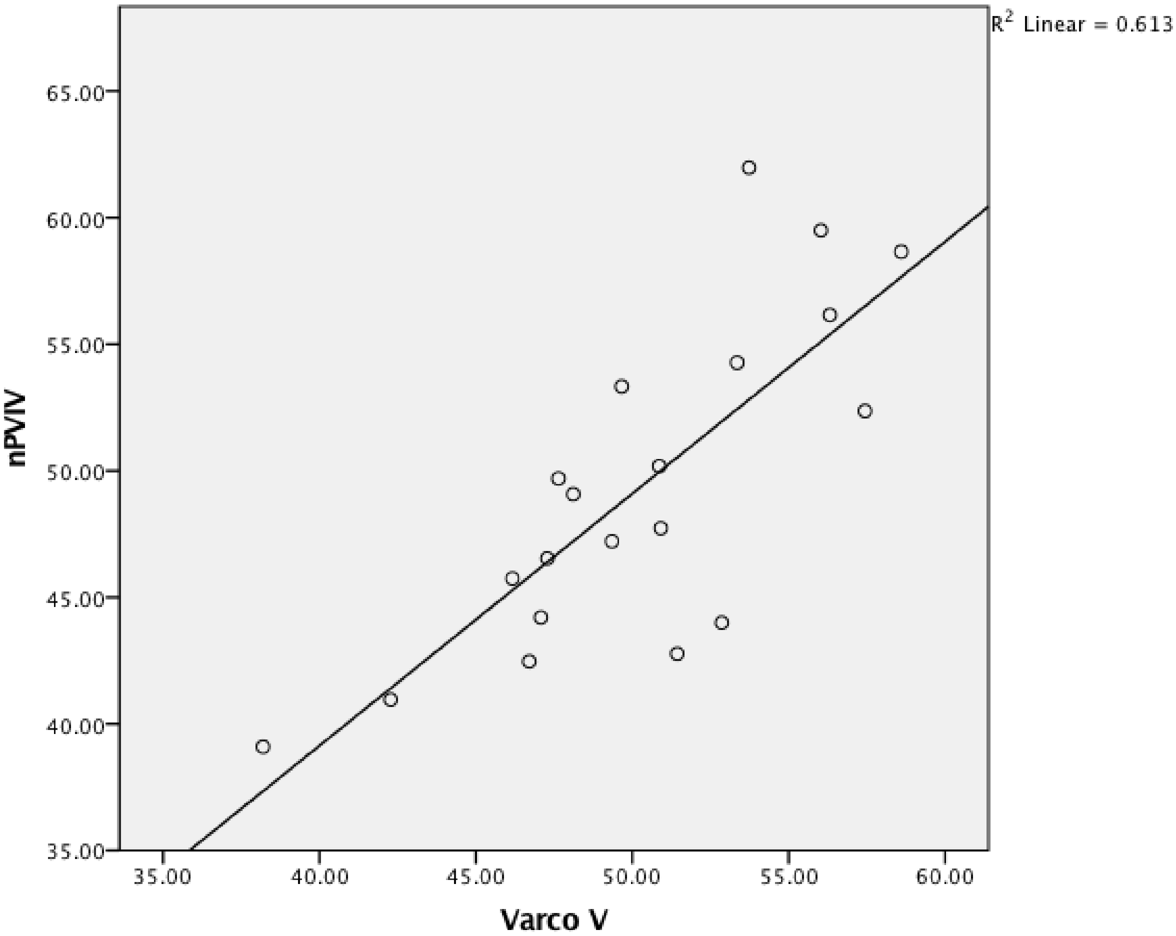

Figure 2. *Scatterplot for nPVI-V and Varco V.*

The second research question is whether the Standard 1, bilingual children aged 6 years, have a different pattern of speech rhythm timing in Setswana in comparison with the children in Standard 2, aged 7 years, who will have had increased exposure to English. The independent samples *t*-test showed a non-statistically significant effect, for both the nPVI-V and Varco V. The nPVI-V results are $t(8) = 283$, $p = .784$. The Varco V results are $t(8) = 537$, $p = .606$. Therefore, the STD 1 group and the STD 2 group's nPVI-V and Varco V are not statistically significantly different. This suggests that the age and standard the children are in did not have an effect on their Setswana speech rhythm.

Discussion

Similar to previous studies (Bunta & Ingram 2007; Grabe et al., 1999; Lleo et al., 2007; Mok, 2011; Whitworth, 2002), the findings of the present study show that there is dissimilarity in the monolinguals' and the bilinguals' development of speech rhythm as shown by the nPVI and Varco V durational measurements. Although the results of the present study support the findings of previous research regarding children keeping the speech rhythm of their two languages distinct (Bunta & Ingram, 2007), the findings themselves are not in the direction anticipated. This is because the Setswana-English bilinguals' speech rhythm is more syllable-timed than that of Setswana monolinguals, as evidenced by the lower nPVI and Varco V means for the bilingual group; that is, they had lower durational variability of syllables in comparison with the monolinguals. It could be said that, in so doing, the bilinguals kept the rhythm of Setswana, which is considered syllable-timed; however, their more syllable-timed Setswana compared to the monolinguals raises questions. At the age of 6-7 years the monolingual children in the present study are old enough to display native-like Setswana rhythm. Therefore, we can expect them to have native-like Setswana rhythm. This expectation is based on the findings of Mok (2011), who established that monolingual Cantonese children and monolingual English children who were 3 years of age showed distinct rhythm patterns in their respective languages. Moreover, the monolingual and bilingual children in Bunta and Ingram (2007), who were around 5 years of age and below, were able to separate the rhythm of their two languages.

Previous research has shown that, at the early stages of speech rhythm development, children's rhythm compared to that of adults is syllable-timed regardless of the rhythm of their language (Allen & Hawkins, 1980; Grabe et al., 1999; Mok, 2011; Ordin & Polyanskaya, 2014, 2015; Payne et al., 2012). Stress timing is acquired later because the children have to learn to reduce syllables with full vowels (Grabe et al., 1999). This could be one of the reasons why the Setswana rhythm of bilinguals is not stressed-timed even though they are dominant in English and have had greater exposure to a British English model, which is considered to be a stress-timed language. The findings of the present study are in support of previous studies on the notion that the speech rhythm of children develops from a low durational variability syllable-timed rhythm because a lower vocalic variability in rhythm timing is easier than one with more variability (Allen & Hawkins, 1980; Grabe et al., 1999; Kehoe et al., 2011; Mok, 2011; Ordin & Polyanskaya, 2014, 2015; Payne et al., 2012). The nPVI and Varco V of the Setswana-English bilinguals produced lower durational variability in the Setswana speech rhythm compared with the Setswana monolinguals. Based on these results, it is possible that the Setswana-English bilinguals who are 6-7 years old in the present study are still at an early stage of Setswana rhythm development, that is, they could be exhibiting incomplete or delayed acquisition of their L1 (Setswana). Most of the children in the previous studies (Bunta & Ingram, 2007; Kehoe et al., 2011; Lleo et al., 2007; Mok, 2011) are younger than the children in the present study. Therefore, the expectation is at the age of 6-7 years the bilingual children's speech rhythm should be fully developed to produce a statistically similar result to that of monolinguals of the same age, particularly for a linguistically less marked language considered syllable-timed like Setswana – and especially in view of the fact that all the children selected for this study spoke Setswana as their first language, had never lived outside Botswana and were introduced to English at the age of 3 years when they started nursery school. Similar to Kehoe et al. (2011), the findings support the idea that the language of the larger community does not necessarily contribute to the rhythm pattern of the bilinguals.

In addition, a number of studies have shown that variability in children's speech increases with age (Bunta & Ingram, 2007; Ordin & Polyanskaya, 2014; Payne et al., 2012). Bunta and

Ingram's (2007) monolingual English adults' nPVI-V scores were higher than that of older children (3.9-5.2 years old) while those of older children were higher than that of the younger children. Therefore, the 6-7-year-old bilingual children in this study should display higher durational variability similar to that of their monolingual peers. The findings of the present study could mean acquisition of high durational variability rhythm by bilingual children similar to that of monolinguals is after the age of 7 years. This supports Whitworth (2002), who noted that complete acquisition of speech rhythm by bilinguals was not evident until around the age of 11 years. It is noteworthy that the data of the present study cannot ascertain if the Setswana-English bilinguals' underdeveloped speech rhythm is due to acquisition delay or incomplete acquisition, because there is nothing in the literature that indicates when the phonology of Setswana is acquired.

While the findings of the present study support the position in other research, that the learning of any language whether L1 or L2 develops from a lower durational variability towards a higher durational variability (Ordin & Polyanskaya, 2014), the Setswana-English bilingual children's low durational variability compared to monolinguals in the present study demonstrates that this does not only occur when the language being acquired is stress-timed. Even when the target language is an L1 considered syllable timed, learning develops from low durational variability. Previous studies (Bunta & Ingram, 2007; Grabe et al., 1999; Kehoe et al., 2011; Lleo et al., 2007; Mok, 2011; Ordin & Polyanskaya, 2014) reported lower variability only when the language being acquired is 'stress-timed'. When the language is 'syllable-timed' there is no significant difference in the rhythm scores of learners compared to that of advanced speakers, as shown by bilingual speakers in Bunta and Ingram (2007) and Mok (2011). The only study to date, known to the researchers, which showed statistically significant differences between monolinguals' and bilinguals' stressed-timed German rhythm was Kehoe and Lleo (2005). The findings of the present study suggest that acquisition of speech rhythm by bilinguals similar to that of monolinguals of the same age might not necessarily depend on the rhythm of the language under investigation (whether syllable-timed or stress-timed).

It is worth noting that there were individual differences, as some of the bilingual participants' rhythm tended towards a high durational variability, though this was not statistically significant. It was also only noticed with the Varco V scores and not with nPVI. Mok (2011) and Low et al. (2002) state that metrics that measure durational variability globally (whole utterance) such as Varco V usually have higher scores than metrics which measure durational variability locally (between successive vowels) like the nPVI. The observed individual differences could, therefore, be attributed to the metric applied. There could also be differences in exposure to Setswana.

Even though the present study could not ascertain what could have taken place in the development of the bilinguals' Setswana rhythm, it is probable that language dominance has played a major role in the different ways in which rhythm patterns in the bilinguals and monolinguals speech developed. Language dominance is often determined by computing Mean Length of Utterance (MLU) for each language the bilingual speaks (Yip & Matthews, 2000). MLU is the number of morphemes or words in a child's intelligible spontaneous utterance. Since the focus of the present study is on phonology rather than morphology, MLU was not measured. However, in the language background questionnaire information provided by the parents, and through the interaction the researcher had with the participants, it was apparent that English was the dominant language of the bilingual group. The results of the present study demonstrate that language dominance has an effect on the development of the bilinguals' speech rhythm, even though that effect was not in the expected direction.

Since exposure to high English input has an effect in the speech of the Setswana-English bilinguals compared to Setswana monolinguals, it was necessary to determine if increased levels of exposure has had an effect in the speech rhythm of 7 years old STD two Setswana-English bilinguals compared to 6 years old STD one Setswana-English bilinguals. The rationale for this comparison was that being a year older and in a higher grade means that this group of bilinguals have been exposed to English for longer. The results did not produce a statistically significant difference. This could mean that the English of children in STD 1 is already at a point where it has had an effect on their Setswana rhythm. The increased levels of exposure to English in the Setswana-English bilinguals is only noticeable when the nPVI and Varco V scores of the bilinguals are compared to those of their monolingual peers, not within the bilingual group.

Conclusion

The present study has demonstrated that, among Setswana-English bilingual children, the rhythmic type of the target language (Setswana), the age of the participants, and the language environment of the bilinguals' upbringing, which involves a high level of English language input, may not necessarily have a strong effect on the developing speech rhythm of those bilinguals. The findings of the present study demonstrate, however, that linguistic input, which led to language dominance, is crucial in the acquisition of speech rhythm. While the dominant language of the larger community is Setswana, for the bilingual group, English is the dominant language because of the high English input they are exposed to at school, at home with their parents and with their friends. As such, their everyday communication is mostly carried out in English. British English, which is stress-timed, is a dominant model in classroom materials; however, although the rhythm of the bilinguals in this study is somewhat dissimilar to their monolingual peers, they do not show British English stress-timing patterns in their Setswana. The suggestion here is that this is because they are still in the process of developing native-like Setswana speech rhythm, and that the introduction of English causes differences in the development of that rhythm in comparison with their monolingual peers. This study also suggests it is because the development of speech rhythm in any language begins with a low vocalic variability (Bunta & Ingram, 2007; Ordin & Polyanskaya, 2014; Payne et al., 2012) that English stress-timing has probably not influenced the Setswana speech rhythm of the bilinguals .

What might throw further light on the findings we present here is the English speech rhythm of the children in this study and of the speakers of Botswana English in their social circle. As English was not the focus of the study, we did not collect contemporaneous English speech data. Future research is, therefore, planned to address this situation. It would also be interesting to investigate the patterns of Setswana and English speech rhythm in the Batswana adult population and compare the findings presented here, as currently there is no empirical research available; it is possible that Botswana English as a global variety is less stress-timed than, e.g., British English, and that the bilingual children were exposed to this variety more than classroom models based on British or American native-like patterns.

References

Abercrombie, D. (1967). *Elements of general phonetics* (Vol. 203): Edinburgh University Press Edinburgh.

Allen, G. D., & Hawkins, S. (1980). Phonological rhythm: Definition and development. *Child phonology, 1*, 227-256.

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*(3), 351-373.

Bayram F., Wright C. (2018) Turkish Heritage Language Acquisition and Maintenance in Germany. In: Trifonas P., Aravossitas T. (eds) Handbook of Research and Practice in Heritage Language Education. Springer International Handbooks of Education. Springer, Cham.

Boersma, P., & Weenink, D. (2007). Praat: doing phonetics by computer (Version 4.6. 09)[Computer program]. Retrieved June 24, 2007.

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.

Coetzee, A. W., & Wissing, D. P. (2007). Global and local durational properties in three varieties of South African English. *The Linguistic Review, 24*(2-3), 263-289.

Cole, D. T. (1955). *An introduction to Tswana grammar*: London Longman.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics 11*(1), 51-62.

Flege, J.E. (2002). Interactions between native and second-language phonetic systems. In P. Burmeister, T. Piske & A. Rohde (Eds.), An intergrated view of alanguage development: papers in honor of Henning Wode (pp.217-244). Trier: Wissenschaftlicher Verlag.

Fuchs R. (2016) Speech Rhythm in Varieties of English. In: Speech Rhythm in Varieties of English. Prosody, Phonology and Phonetics. Springer, Berlin, Heidelberg.

Gibbon, D., & Gut, U. (2001). *Measuring speech rhythm.* In *Eurospeech, 2001*, 95-98.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology, 7*(515-546).

Grabe, E., Post, B., & Watson, I. (1999). *The acquisition of rhythmic patterns in English and French.* Paper presented at the Proc. ICPhS.

Gut , U. (2012). Rhythm in L2 speech . In D. Gibbon , D. Hirst , & N. Campbell(Eds.), Speech and language technology: Volume 14/15, special edition. Rhythm, melody and harmony in speech. Studies in honour of Wiktor Jassem (pp. 83 – 94). Pozna ń, Poland : Polskie Towarzystwo Fonetyczne .

Kehoe. (2015). Lexical-phonological interactions in bilingual children. *First Language, 35*(2), 93-125.

Kehoe, Lleó, C., & Rakow, M. (2011). Speech Rhythm in the Pronunciation of German and Spanish Monolingual and German-Spanish Bilingual 3-Year-Olds. *Linguistische Berichte, 2011*(227), 323-352.

Kehoe, & Stoel-Gammon, C. (2001). Development of syllable structure in English-speaking children with particular reference to rhymes. *Journal of child language, 28*(2), 393-432.

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, 2011*(227), 323-352.

Knight, R.-A. (2011). Assessing the temporal reliability of rhythm metrics. *Journal of the International Phonetic Association, 41*(3), 271-281.

Li, A., & Post, B. (2014). L2 acquisition of prosodic properties of speech rhythm. *Journal of the Studies in Second language acquisition, 36*(2), 223–255.

Lleó, C., & Kehoe, M. (2002). On the interaction of phonological systems in child bilingual acquisition. *International Journal of Bilingualism, 6*(3), 233-237.

Lleó, C., Rakow, M., & Kehoe, M. (2007). Acquiring rhythmically different languages in a bilingual context. In J. Trouvain & W. J. Barry (eds.), Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS), pp. 1545–1548

Mayer, M. (1969). *Frog, where are you?* : Dial Press: New York.

Mok, P. P. (2011). The acquisition of speech rhythm by three-year-old bilingual and monolingual children: Cantonese and English. *Bilingualism: Language and Cognition*, *14*(04), 458-472.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, *24*(3), 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), 1-19.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication, 41*(1), 233-243.

Nolan, F., & Asu, E. L. (2009). The pairwise variability index and coexisting rhythms in language. *Phonetica, 66*(1-2), 64-77. .

Ordin, M., & Polyanskaya, L. (2014). Development of timing patterns in first and second languages. *System, 42*, 244-257.

Ordin, M., & Polyanskaya, L. (2015a). Perception of speech rhythm in second language: The case of rhythmically similar L1 and L2. *Frontiers in Psychology, 6,* 1–15.

Payne, E., Post, B., Astruc, L., Prieto, P., & Vanrell, M. d. M. (2012). Measuring child rhythm. *Language and Speech*, *55*(2), 203-229.

Pike, K. L. (1945). The intonation of American English. Ann Arbor: university of Michigan Press.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265-292.

Raven, J. (1982). *Coloured Progressive Matrices*. London: Harcourt Publishers.

Reilly, J., Losh, M., Bellugi, U., & Wulfeck, B. (2004). "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and language*, *88*(2), 229-247.

Roach, P. (1982) On the distinction between 'stress-timed' and 'syllable-timed' languages. *In Linguistic Controversies* (D. Crystal, Editor), pp. 73-79, London: Arnold.

Sebina, B., Setter., J. & Wright, C. (2019). The penultimate syllable vowel length among Setswana-English bilingual children. In *Proceedings of the International Congress of Phonetic Sciences* (ICPHS 2019), Melbourne, Australia, 5-9 August 2019.

Setter, J.; Stojanovik, V.; van Ewijk, L.; Moreland, M. Pitch range and vowel duration in the speech of children with Williams Syndrome. *Proceedings of the 16th International Congress of Phonetic Sciences* (ICPhS '07); Saarbrücken, Germany. 2007. p. 1337-1340.

Tan, R. S. K., & Low, E.-L. (2014). Rhythmic Patterning in Malaysian and Singapore English. *Language and Speech*, *57*(2), 196-214.

Turk, & Shattuck-Hufnagel. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong. *Laboratory Phonology, 4*(1), 93-118.

Turk, & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics, 35*(4), 445-472.

White, L., & Mattys, S. L. (2007a). Calibrating rhythm: First language and second language studies. *Journal of Phonetics, 35*(4), 501-522.

White, L., & Mattys, S. L. (2007b). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaro´, & M.-J. Sole´ (Eds.), Segmental and prosodic issues in romance phonology (Current issues in linguistic theory series). Amsterdam, Philadelphia: John Benjamins.

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. (2000). Syntactic transfer in a Cantonese–English bilingual child. *Bilingualism: Language and Cognition, 3*(3), 193-208.