Pertanika J. Soc. Sci. & Hum. 28 (4): 3125 - 3143 (2020)

PERTANIKA JOURNALS

SOCIAL SCIENCES & HUMANITIES

Journal homepage: http://www.pertanika.upm.edu.my/

Performance in Nonword Repetition Tasks among Mandarin-English Bilingual Children in Malaysia

Chai Ping Woon[1], Ngee Thai Yap[1]* and Hui Woan Lim[2]

[1]Department of English, Faculty of Modern Languages and Communication, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[2]Centre for Rehabilitation and Special Needs Studies, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, 50300 Kuala Lumpur, Malaysia

ABSTRACT

The nonword repetition (NWR) task has been used to measure children's expressive language skills, and it has been argued to have potential as an early language delay/impairment detection tool as the NWR task can be conducted rather easily and quickly to obtain a quantitative as well as a qualitative measure of children's attention to lexical and phonological information. This paper reports the performance of two NWR tasks among thirty bilingual Mandarin-English preschoolers between the age of four through six. The study indicated that performance in the NWR tasks showed a developmental trend with older children performing better than younger children. Word length also had a significant effect on performance, possibly an effect from better short-term memory capacity as the child grew older. The children also performed better in the Mandarin NWR task compared to the English NWR task. These findings suggest potential clinical applications for diagnosis of children with language impairment or at risk of language development delay. However, further studies should improve on the tasks to verify its efficacy and to obtain norms for performance with a larger sample of children at various age groups.

Keywords: Bilingualism, language development, Mandarin-English bilingual children, nonword repetition, simultaneous bilinguals

ARTICLE INFO

Article history:
Received: 24 February 2020
Accepted: 09 June 2020
Published: 25 December 2020

DOI: https://doi.org/10.47836/pjssh.28.4.35

E-mail addresses:
reginawoon@hotmail.com (Chai Ping Woon)
ntyap@upm.edu.my (Ngee Thai Yap)
lim@ukm.edu.my (Hui Woan Lim)
* Corresponding author

INTRODUCTION

The nonword repetition (NWR) task has been used to measure children's expressive language skills in normal and abnormal language development. Researchers often use the NWR task to study the mechanisms

ISSN: 0128-7702
e-ISSN 2231-8534

of phonological short-term memory (STM) underlying children's language development. The NWR task has been recognized as a potential psycholinguistic tool to identify children with specific language impairment (SLI). SLI is considered a neuro-developmental disorder that affects language development among children; but diagnosis of these children remains challenging. Children with SLI show language ability below the language skills expected for their age, but they have no hearing impairment, neurological damage nor motor problems that affect speech production (Leonard, 2000). In addition, they also do not exhibit any clinical indicators of autism, and they have normal intellectual abilities (Leonard, 2000). The protocol followed by speech language pathologist to identify SLI children involves a series of tests and the language assessment tool is only one of the tests used.

The process of identifying children at risk of SLI at an early age is challenging even among monolingual children. Often, it is difficult to differentiate normal developmental delay from SLI which implicates language impairment. The problem is even more challenging in Malaysia as many Malaysian children are brought up in a bilingual or multilingual environment from a very young age (Lim, 2018; Lim et al., 2015). Most assessment tools that are used are either not normed or normed mainly on monolingual children and may not be suitable for identifying SLI among bilingual children. (Lim, 2018; Lim et al., 2015). Currently, there are only three published child grammar tests for the Malay and Chinese children (A Razak et al., 2016; A Razak et al., 2014; A Razak et al., 2018). These language assessment tools are language dependent and there has yet been a study exploring the use of the nonword repetition task to compare performance of simultaneous bilingual children in the two languages in their repertoire. Devising language assessment tools that are language-specific such as Mandarin and English are highly desirable since professionals (e.g. speech-language pathologists) dealing with the Malaysian children are from diverse language backgrounds (Ibrahim et al., 2019).

Malaysia is a multilingual country with most Malaysians speaking at least two languages but many particularly the ethnically Chinese may speak as many as five languages (see discussion in Yap et al., 2017). As discussed in Lim et al. (2015) and Lim (2018), Malaysian Chinese children typically receive regular exposure to Mandarin and English from birth, at home, in the kindergarten and from the community at large (e.g. through interactions with neighbours, at the playground, or shopping complex). They also receive robust exposure to Malay once they enter the kindergarten sometimes from as early as 2-3 years old. Therefore, strictly speaking, the focus of this study is on simultaneous Mandarin-English bilingual children who have started to learn Malay once they enrol in the kindergarten. But for this study, we did not focus on the development of Malay since it is the first study conducted using the NWR task in

the Malaysian context, and the children in this study are likely to be more dominant in Mandarin and English compared to Malay.

Since Malaysia is so linguistically diverse, the potential use of the NWR task as an assessment tool is an attractive option as it is not bound to vocabulary knowledge in the specific language. However, before such a tool can be developed, an understanding of the performance of such tasks with typically developing children is needed to provide baseline information to interpret the results of this task with children exhibiting atypical development. Furthermore, there are no substantial language research on simultaneous Mandarin-English bilingual children in Malaysia. Hence, in this paper, we present results from a study conducted with Mandarin-English bilingual children to examine the feasibility of the NWR tasks as potential language assessment tools.

Nonword Repetition Tasks

The NWR task have been used in the past twenty years by psychologists and researchers to study the mechanism of phonological short-term memory (STM). The task requires participants to listen to a nonword and to repeat it verbatim. The NWR task taps a range of perceptual, cognitive, and motor processes, which are widely acknowledged. Evidence from monolingual and cross-linguistic studies on performance of NWR tasks are related to linguistic knowledge (Archibald & Gathercole, 2006; Chiat & Roy, 2007; Gathercole et al., 1991; Graf Estes et al., 2007; Messer et al., 2010; Vitevitch & Luce, 2005), vocabulary knowledge (Chiat & Roy, 2007; Gathercole & Baddeley, 1989; Gathercole et al., 1999; Hoff et al., 2008; Masoura & Gathercole, 1999, 2005; Summers et al., 2010), and grammatical knowledge (Dispaldro et al., 2011; French & O'Brien, 2008). As argued in these studies, NWR is linked to language learning; hence typically developing children are able to repeat nonwords more accurately compared to those with some form of developmental delay or language impairment. Thus, it is reasonable to believe that NWR has the potential to identify children with language impairment.

NWR task has been widely proposed in English as a potential psycholinguistic assessment tool to detect SLI children from their typically developing (TD) peers (Archibald, 2008; Bishop, 2004; Coady & Evans, 2008; Conti-Ramsden et al., 2001; Dollaghan & Campbell, 1998; Ellis-Weismer et al., 2000; Gathercole & Baddeley, 1990; Graf Estes et al., 2007). These studies found that SLI children were extremely weak in the NWR task, possibly due to impairment or a capacity limitation of the phonological component of working memory. According to Baddeley (2003), working memory (WM) is a cognitive system that involves a combination of storage and cognitive processes and operations involving linguistic information. The multi-component model of WM developed by Baddeley explained how memory and linguistic knowledge influence the repetition task. Baddeley's model of the WM is comprised of four components: 1) phonological loop, 2)

visuospatial sketchpad, 3) central executive, and the 4) episodic buffer. The phonological loop comprises a phonological store, which can retain information for a few seconds before it decays and an articulatory rehearsal system, which can refresh and rehearse the information. The visuospatial sketchpad is responsible to maintain visual and spatial information for a short period. The central executive contains a set of cognitive processes that interact with other components and long-term memory (LTM). The episodic buffer behaves like a backup store, with a limited capacity capable of reinforcing the phonological loop or the visuospatial sketchpad, and to integrate information from many different sources. It also links STM to LTM.

A meta-analysis of studies investigating different NWR task performance reported across different studies between children with or without SLI carried out by Graf Estes et al. (2007), showed that children with SLI performed significantly lower than children without SLI on longer nonwords (3- to 4-syllable nonwords) than shorter nonwords (1- to 2- syllable nonwords). Gathercole and Baddeley (1990) viewed the repetition difficulties as a reflection of the capacity limitation of the phonological component in the working memory. Therefore, a deficit in the capacity of phonological STM may lead to difficulty in repeating nonwords with more syllables. The relationship between poor performance in the NWR task and language impairment has also been found across different languages: Spanish (Girbau & Schwartz, 2007; Windsor et al., 2010), Italian (Bortolini et al., 2010), Dutch (De Bree et al., 2007; Rispens & Parigger, 2010), French (Thordardottir & Brandeker, 2013), Japanese (Kosaka, 2009), and Mandarin (Chi, 2007). Chi (2007), who investigated phonological memory on NWR performance among Taiwanese Mandarin-speaking children with SLI found that the children with SLI in the study performed poorly in two-, three-, four-, and five-syllable nonwords compared to their TD peers. She concluded that the children with SLI in the study did in fact have difficulties in the NWR task.

Cross-linguistic studies of the performance on NWR tasks and word learning among bilingual children from different linguistic backgrounds have been reported to rule in or rule out bilingual children with or without language impairment. Girbau and Schwartz (2007) also used the NWR task with bilingual Spanish-English children with and without SLI. The result showed that SLI children performed lower in the NWR tasks compared to their TD peers. Thordardottir (2017) found that second language (L2) learners of Icelandic language could also obtain high NWR scores, which was important in terms of the use of NWR tests for the purpose of ruling in or ruling out language impairment in this population. In contrast, only one study (Stokes et al., 2006) had reported findings that argued against the potential utility of NWR tasks. Their results demonstrated that Cantonese-speaking children with SLI did not perform significantly lower compared to their TD peers. They further argued that

when only phonological working memory skills were tested, the SLI children's ability to repeat nonwords did not differ from the TD children. It is evident that more studies are needed to determine the efficacy of the NWR task as a diagnostic tool to help identify children with language development delay or language impairment such as SLI.

In the context of Malaysian children, as mentioned earlier, there has been no substantial studies conducted to document the language development milestones of simultaneous Mandarin-English bilingual children. Lim et al. (2015) and Lim (2018) examined the error patterns in terms of phonological development of multilingual Malaysian children in three languages: Malay, Mandarin and English using real words that were either monosyllabic, disyllabic or trisyllabic as the target group for that study was 2-4 year old children. Therefore, there is a need to examine the language developmental milestones for simultaneous bilingual children in the preschool age range and to examine the effect of word length in the novel nonword repetition tasks. The study reported in this paper was part of a larger study conducted to examine how typically developing simultaneous Mandarin-English bilingual children in Malaysia would perform in NWR and sentence repetition tasks conducted in English and Mandarin across the different age groups. However, due to space limitation, this paper will report only the findings from the NWR tasks (see Woon (2015) and Woon et al. (2014) for a discussion of the sentence repetition tasks). Specifically, this paper addresses the following questions: 1) Is there a significant difference in the performance of Mandarin-English bilingual children across the different age groups on the Mandarin and English NWR tasks? 2) Is there a word length effect in the performance of the Mandarin and English NWR tasks across the different age groups? and 3) Is there a significant difference between the performance of the children in the Mandarin NWR task compared to the English NWR task?

METHOD

Participants of the Study

The participants were selected using purposive sampling with specific inclusion and exclusion criteria. As the study is interested in simultaneous bilingual children, only children who fit the above criteria were recruited. Information about the children's ability in both languages were determined by the report provided by their teachers which were available to the principal of the kindergarten. The principal of the kindergarten selected the children to take part in this study based on the teacher's recommendation which were made based on the children's assessment at the kindergarten. Children were only included if they had not been referred for speech and language therapy, and they had no hearing loss, no history of neurological impairment, developmental disorders or social, emotional or behavioral difficulties as reported by their parents or teachers.

The children were recruited from a kindergarten in Kuala Lumpur. The kindergarten was selected from among the kindergartens in the Klang valley as it was a large kindergarten with about 5 classes for each age group. The children from the kindergarten had regular exposure to the two target languages: English and Mandarin since they had started attending nurseries from the age of 2;00 to 3;00. The children were dominant in both English and Mandarin as reported by the teachers and parents and as observed by the researcher who conducted the tasks. Research information, consent forms, and a parental questionnaire were distributed to the parents whose children were selected for the study; informed consent were obtained prior to the data collection sessions.

Thirty typically developing (TD) children from ages 3;9 to 6;5 years were recruited for the study. Fourteen of the children were boys and sixteen were girls. For the purpose of data analysis, the children were divided into three groups according to their biological age: 4-year-olds, 5-year-olds, and 6-year-olds. Table 1 presents the demographic information of the participants.

All the participants were ethnically Chinese. They all spoke Mandarin and English. All children were reported to have early exposure to Mandarin and English before the age of three. They are early simultaneous Mandarin-English bilingual children following the lenient cut-off age criteria, namely regular exposure to two or more languages before three (Lim, 2018; Lim et al., 2015). Information about their language use and family background was obtained from a parental and teacher questionnaire adopted from Gutiérrez-Clellen and Kreiter (2003).

English Nonword Repetition Task

The Children's Test of Nonword Repetition (CNRep), developed by Gathercole and Baddeley (1996) was adapted for use in this study. The CNRep consists of 40 nonwords, from two to five syllables (e.g. /'hɑmpənt/, /'brɑstərə/, /stɒpə'grɑtɪk/, /prɪstər'ɑkʃənl/). Half of the nonwords contain syllables with either initial or final consonant clusters (e.g., /bl/, /nt/), and the other half contain syllables with simple onsets and codas. The CNRep contains high word-like nonwords with existing English derivational morphemes, such as, -ing in /'glIstər**ɪŋ**/, -ist in /kəntrampən**ist**/, -ic in /wʊgəlæm**ɪk**/. After the pilot study, which was conducted with six children, two children from each age group, the five syllable nonwords were omitted. The children in the pilot study

Table 1
Demographic information of the participants

Group	*N*	Age range (year;month)	Gender (boys/girls)	Mean age (month)	*SD*
1	10	3;9 – 4;6	6/4	49.7	2.63
2	10	4;7 – 5;5	4/6	60.1	3.93
3	10	5;9 – 6;5	4/6	73.7	3.16
Total	30	-	14/16	-	-

were found to have difficulty in repeating five-syllable nonwords and they lacked confidence and patience. The final version of the task consisted of only 30 nonwords. The English nonword stimuli were recorded by a Malaysian English teacher. The stimulus items were read with the usual Malaysian accent to ensure that the nonwords did not sound too foreign to the children.

Mandarin Nonword Repetition Task

The Mandarin Nonword Repetition (MNWR) Test that was developed by Chi (2007), was adapted for use in this study. The test incorporated Mandarin lexical knowledge in the same way done in CNRep to ensure that the test items created were high word-like nonwords. A word in Standard Mandarin is a set of monosyllables, represented by one or a combination of monosyllables. The original MNWR test from Chi (2007) contains a total of 55 nonwords: 10 two-syllable nonwords (e.g. gu1tong2 姑同), 11 three-syllable nonwords (e.g. ge2ba1tian2 格八甜), 13 four-syllable nonwords (e.g. ben4kui4rou2bao3 笨愧柔跑), 8 five-syllable nonwords (e.g. pei4tuan2ben3zuo4jiu1 佩团本作纠), and 3 six-syllable nonwords (e.g. huai2zhi4shu1jiao1li3te4 怀直书交李特). In order to match the ENWR task which had only 30 nonwords, the five- and six-syllable nonwords were omitted from this study. In addition to the selection of words, the nonwords were also selected to match the number of items for each word length found in the English test.

The Mandarin nonwords were recorded by the first author who is also a Mandarin-English bilingual speaker. The recordings were done using a Sony ICD-UX513F Digital Voice Recorder. The MNWR task was also piloted with the same six children. None of these children were recruited for the actual study as there were substantial differences in the NWR task as well as the sentence repetition task that were conducted.

Procedure

The participants were tested individually in a quiet room at their kindergarten. All sessions were audio-recorded. Before each test, the children were given two examples to practice with. The children were told that they would listen to some made-up words, and were asked to repeat the words exactly the way they heard them. Each word was presented once, followed by a three second pause during which the participants were required to repeat the word heard. The children were allowed to listen to each nonword only once. When a child was not able to repeat a word within the three seconds, the audio file was stopped and he or she was given time to make a response. The children were praised for any attempt at repetition. The participants were required to wear a microphone headset. No reply and live-voice was presented by the examiner. The audio files were presented on a laptop computer in a fixed random order from two to four syllables. The participants' responses were recorded using the same voice recorder mentioned earlier. The NWR tasks in each language took around three minutes to complete for each child.

Scoring

The whole word scoring method was adopted in this study. The nonwords were scored offline with each correct response awarded with one point and no points were awarded for an incorrect production. The nonwords were pronounced correctly when there was no omission, substitution or addition of phonemes. The phonetic variation of a particular phoneme was not counted as an error if that variation was not pronounced like any other phonemes. Only a few instances of tone errors were found in the MNWR task as also shown in Lim (2018) who found that tone errors were rather rare. The only instances of error found in the study were the change from the falling-rising tone to the rising tone, for example, zao3 (早) to zao2 which can be regarded as a slip of the tongue.

Inter-Rater Reliability for Scoring and Transcription

Audio-recorded responses from the participants were scored and transcribed by a second rater, who was a postgraduate student and a Mandarin-English bilingual speaker. After comparing the scores and transcriptions, both raters discussed and resolved any discrepancies by listening again to the audio-recorded responses. The Cohen's kappa (k) statistic was used to measure the inter-rater reliability score and the agreement between the two raters was high (ENWR: $k = 0.926, p < 0.001$; MNWR: $k = 0.963, p < 0.001$).

Data Analysis

As the sample size for each age group in the study was small, the non-parametric tests were used in the analyses. Within language group differences for overall task accuracy and word length effects on the NWR tasks were examined for each language version of the NWR task.

RESULTS

To answer the first research question about group differences in the NWR tasks, the group differences were examined using the overall task accuracy for each NWR task. Table 2 presents the performance of participants in the various age groups. The results show that the older children performed better than the younger children.

The Kruskal-Wallis test was conducted to evaluate the difference among the three age groups (ages 4, 5, and 6) on median change in the performance of the English nonword repetition task (ENWR). The results of the analysis indicated that there was a significant difference in performance

Table 2
Performance of Nonword Repetition Tasks by age

NWR Task	4-year-old		5-year-old		6-year-old	
	Mean	***SD***	**Mean**	***SD***	**Mean**	***SD***
Mandarin	11.1	3.2	15.6	2.4	15.4	3.7
English	8.8	2.9	11.9	4.0	14.1	3.4

of ENWR between the three age groups, in the medians, $\chi^2(2, N = 30) = 9.02$, $p = 0.011$, with a mean rank of 9.30 for age 4, 16.20 for age 5, and 21.00 for age 6. Follow-up tests were conducted to evaluate pairwise differences among the three groups using the Mann-Whitney *U* test. The results indicated a significant difference only between the 4-year-olds and the 6-year-olds [$U = 11.5$, $p = 0.002$]. The comparison between the 4-year-olds and the 5-year-olds [$U = 26.5$, $p = 0.075$] and between the 5-year-olds and the 6-year-olds [$U = 33.5$, $p = 0.218$] were not statistically significant as shown in Table 3.

With regard to the Mandarin nonword repetition task (MNWR), there was a significant difference between the median of the groups [$\chi^2(2, N = 30) = 9.91$, $p = 0.007$] with a mean rank of 8.40 for age 4, 19.55 for age 5, and 18.55 for age 6. The pair wise comparison using the Mann-Whitney Test as shown in Table 4 also found a significant difference only between the 4-year-olds and the 5-year-olds [$U = 12.0$, $p = 0.003$] and between the 4-year-olds and the 6-year-olds [$U = 17.0$, $p = 0.011$]. The difference in the performances between the 5-year-olds and the 6-year-olds was not statistically significant [$U = 47.5$, $p = 0.853$].

To answer the second question, the group differences in NWR tasks of both languages were examined for overall task accuracy for each word length. A Friedman test was conducted to evaluate the differences in median among the children's performance with different word lengths in the ENWR and MNWR. Table 5 presents the means, standard derivation, and median for different age groups and different word lengths for the ENWR task.

The results indicated that there was a significant difference among the three group with different word length [age 4,

Table 3
Pairwise comparison for the ENWR task by age group

	Age 4 – Age 5	**Age 4 – Age 6**	**Age 5 – Age 6**
Mann-Whitney U	26.500	11.500	33.500
Wilconxon W	81.500	66.500	88.000
Z	-1.787	-2.926	-1.260
Asymp.Sig. (2-tailed)	0.074	0.003	0.208
Exact sig. [2*(1-tailed sig.)]	0.075	0.002	0.218

Table 4
Pairwise comparison for the MNWR task by age group

	Age 4 – Age 5	**Age 4 – Age 6**	**Age 5 – Age 6**
Mann-Whitney U	12.000	17.000	47.500
Wilconxon W	67.000	72.000	102.500
Z	-2.884	-2.509	-.191
Asymp.Sig. (2-tailed)	0.004	0.012	0.849
Exact sig. [2*(1-tailed sig.)]	0.003	0.011	0.853

$\chi^2(2, N = 10) = 14.76$, $p = 0.001$; age 5, $\chi^2(2, N = 10) = 14.22$, $p = 0.001$; age 6, $\chi^2(2, N = 10) = 14.80$, $p = 0.001$.] Follow-up pairwise comparisons were conducted using a Wilcoxon signed-rank test with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. The results indicated that there was no significant differences between the two- and three-syllable nonwords [age 4: z =2.360, $p = 0.018$; age 5: $z = -2.200$, $p = 0.028$] or between the three- and four-syllable nonwords among the age four and age five groups [age 4: $z = -2.372$, $p = 0.018$; age 5: $z = -2.336$, $p = 0.019$]. However, there was a significant difference in the performance of the two-syllable and the four-syllable nonwords [age 4: $z = -2.827$, $p = 0.005$; age 5: $z = -2.842$, $p = 0.004$]. For the six-year-olds, the results showed that the children performed significantly better with the two-syllable nonwords compared to the four-syllable nonwords [$z = -2.825$, $p = 0.005$] and between the three- and four-syllable nonwords [$z = -2.555$, $p = 0.011$], but there was no significant differences between the two- and three-syllable nonwords [$z = -1.552$, $p = 0.121$]. Figure 1 presents the mean scores of the subjects for both the English NWR tasks.

Like the ENWR task results, the children's performance for the MNWR task was also influenced by the length of the nonwords. The tests indicated that there was a significant difference between groups with different word lengths [age 4: $\chi^2(2, N$

Table 5
Mean, standard deviation and median of ENWR scores by age and word length

Word length	Age 4			Age 5			Age 6		
	M	*SD*	Median	*M*	*SD*	Median	*M*	*SD*	Median
2 syllables	4.2	1.14	4.5	5.5	1.35	5.5	6.1	1.66	5.5
3 syllables	3.0	1.33	2.5	4.1	1.45	4.5	5.2	1.48	4.5
4 syllables	1.6	1.07	2.0	2.3	2.00	1.5	2.8	1.32	5.0

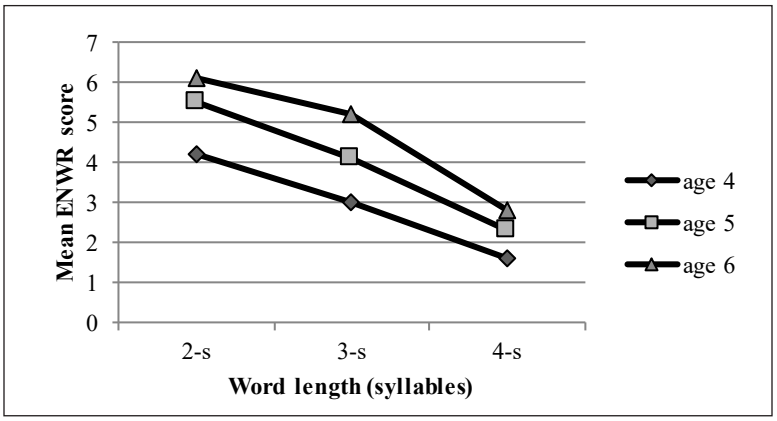

Figure 1. Mean scores for English nonwords by word length and age

= 10) = 10.06, p = 0.007; age 5: $\chi^2(2, N = 10) = 18.20$, $p = 0.000$; age 6: $\chi^2(2, N = 10) = 13.63$, $p = 0.001$]. Follow-up pairwise comparisons were conducted using the Wilcoxon signed-rank test with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. The results indicated that among the age 4 and age 6 groups, there was no significant difference between the two- and three- syllable nonwords [age 4: z = -2.259, p = 0.024; age 6: z = -2.322, p = 0.020] or between the three- and four-syllable nonwords [age 4: z = -1.734, p = 0.083; age 6: z = -2.263, p = 0.024]; however, there was a significant difference when comparisons were made between the two- and four-syllable nonwords [age 4: z = -2.620, p = 0.009; age 6: z = -2.825, p = 0.005]. In contrast, for the 5 year-olds, there was a significant difference between the two- and three-syllable nonwords [z = -2.676, p = 0.007] and between the three- and four-syllable nonwords [z = -2.821, p = 0.005], and also between two- and four-syllable nonwords [z = -2.820, p = 0.005]. Table 6 presents the overall mean scores, standard deviation, and median for the MNWR task. Figure 2 presents the mean of MNWR performance with different word lengths for different age groups.

In order to answer the third question, the overall task accuracy for each language was examined with the Mann-Whitney U test to determine the effects of the language (Mandarin vs English) on the NWR tasks. The results showed that the children performed better in the MNWR task compared to the ENWR task [MNWR: M =

Table 6
Mean, standard deviation and median of MNWR scores by age and word length

Word Length	Age 4			Age 5			Age 6		
	M	*SD*	Median	*M*	*SD*	Median	*M*	*SD*	Median
2 syllables	5.4	1.35	5.0	7.7	0.95	8.0	7.6	1.43	7.5
3 syllables	3.2	1.69	3.0	5.2	1.40	5.0	5.1	2.45	5.5
4 syllables	2.5	1.58	2.0	2.7	1.16	2.5	2.7	1.34	2.5

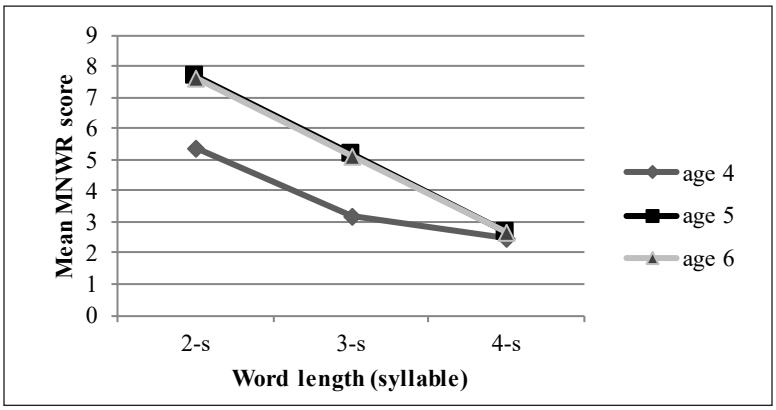

Figure 2. Mean scores of Mandarin nonwords by word length and age

14.03, *SD* = 3.66; ENWR: *M* = 11.60, *SD* = 4.01). Table 7 presents the mean rank and sum of ranks for the Mandarin and English NWR tasks. The analysis of Mann-Whitney test presented in Table 8 shows that there was a significant difference between the performance of these two language tasks. It can be interpreted that the performance in MNWR was better than ENWR ($U = 285.5$, $p = 0.015$). Figure 3 shows the mean scores of the ENWR task and the MNWR task by age groups.

DISCUSSION

Overall, the older children (5- and 6-year-olds) in this study performed better than the four-year-olds, and the children performed better in Mandarin compared to English. The overall results are consistent with earlier findings (e.g. de V. Hage & Grivol, 2009; Gathercole & Baddeley, 1989; Hoff et al., 2008; Stokes et al., 2006) where older children performed better than younger ones. The mean scores in both NWR tasks increased with age of the participants.

Table 7
Performance of NWR in Mandarin and English

NWR tasks	N	Mean rank	Sum of ranks
Mandarin	30	35.98	1079.50
English	30	25.02	750.50

Table 8
Comparison of performance in the Mandarin and English NWR tasks

	MNWR - ENWR
Mann-Whitney U	285.500
Wilcoxon W	750.500
Z	-2.440
Asymp.Sig. (2-tailed)	0.015

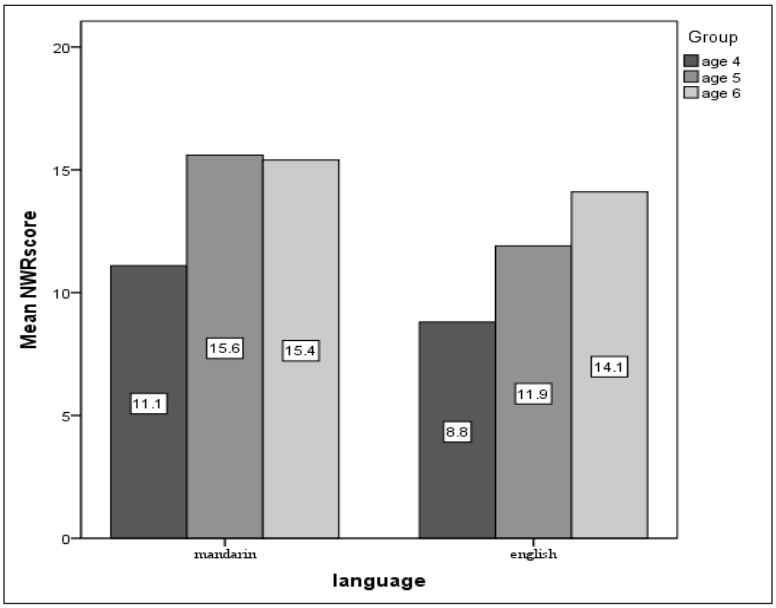

Figure 3. Mean scores of MNWR and ENWR tasks by age

The NWR tasks involved listening to segmental speech sounds presented in a sequence, remembering them, and repeating them. Hulme et al. (1984) found that the development of phonological STM, which increased with age, could be connected to the increased speed and accuracy of recall and it was related to the development of speaking skills. Hoff and her colleagues (2008) explicitly pointed out that children who were more advanced in phonological development had more phonological representations. This could imply that performance in the NWR task relies on phonological STM, as phonological development is indexed by age; the older children are, the better they are expected to perform in remembering new word forms.

The results of this study also support the idea that performance in the NWR task reflects language development of children at different age groups. Although the 4-year-old children performed poorly on the NWR tasks, they had no difficulty in completing the task. This is not the case with the sentence repetition task which requires stronger language skills and knowledge of the grammatical system. Also the performance of the NWR for both languages showed that the children did not reach ceiling level performance even with the 6-year-olds in both language tasks, as compared to the sentence repetition tasks reported in Woon (2015) and Woon et al. (2014). This suggests that the NWR task has better potential in being used as a diagnostic tool since it is a relatively easy task to conduct and it does not require the child to have acquired the grammatical system of the language.

The accuracy scores also revealed that the participants had more difficulties with the longer nonwords than the shorter ones. Earlier studies had suggested that memory span was related to word-length effects (e.g. Archibald, 2008; Baddeley et al., 1975; Gathercole, 2006), and this explained why the longer words were more poorly recalled compared to the shorter words. The phonological loop proposed by Baddeley and Hitch (1974) gives a comprehensive explanation about the word-length effect. Working memory can only hold a phonological form for a very short while. Therefore, performance in the NWR task is expected to be affected by time-based decay effects. Longer nonwords require more time to be presented and repeated. The phonological representation of longer words may decay at a greater extent before the participants could repeat and rehearse it in their mind. Furthermore, the children could not resort to guessing as semantics could not be called upon to assist them in remembering as the stimulus items were all nonwords. Thus, the children's performance in the NWR tasks for both English and Mandarin were affected by the length of the nonwords counted in terms of number of syllables.

Previous studies found that the accuracy rate of NWR was related to word-likeness, phonotactic probability, articulatory complexity, and other linguistic factors (e.g. Archibald & Gathercole, 2006; Chiat & Roy, 2007; Graf Estes et

al., 2007; Messer et al., 2010; Vitevitch & Luce, 2005). In fact, the lack of support of STM or the inability to create a long term memory representation resulting from very limited lexical knowledge in a particular language could lead to difficulty in repeating uncommon phoneme sequences (Archibald, & Gathercole, 2006). Furthermore, Gathercole (2006) suggested that familiarity of constituent segment influenced the accuracy of NWR performance. It is therefore not surprising to find the Malaysian bilingual children performing better in the Mandarin task compared to English task as it may be possible that despite our effort to identify children who were simultaneous bilinguals with exposure to both languages from before the age of three, there could be variability in the sampling which cannot be avoided in the absence of an accurate tool to measure exposure to the two languages.

Apart from the children's phonological STM capacity, the performance of the children in the NWR tasks could also be influenced by inherent differences in the two languages: Mandarin and English. As discussed in Duanmu (2007), there are two important differences between English and Chinese words. Most Chinese words are monosyllabic, while only 13% of English words are monosyllabic; most words in English are polysyllabic. Therefore, it may be possible that phonological STM plays a more prominent role in the learning of English words while semantic knowledge may be used to help children remember longer sequences of compounding available in Chinese words. Thorn and Gathercole (1999) argued that phonological STM functioned in a highly language-specific way. This could possible explain why the performance of the children in the Mandarin NWR task reached a plateau and there was no difference between the performance of the 5-year old and the 6-year olds for all categories of word-length. Nevertheless, this remains an open question for now that needs to be addressed in future studies as the removal of the 5 and 6-syllable nonwords could have impacted on the findings.

CONCLUSION AND RECOMMENDATIONS

The previous discussion led to the following conclusions. First, the NWR tasks can be used to assess language development among the Malaysian Mandarin-English speaking children. The tasks show that the MNWR task was able to distinguish the four-year-olds from the five- and six-year-olds, but it was not adequately sensitive to distinguish the five-year-olds from the six- year-olds. The ENWR task, on the other hand was only able to distinguish the four-year-olds from the six-year-olds. The results also suggest that the NWR tasks are challenging as the older children did not reach the ceiling level unlike the sentence repetition tasks reported in Woon (2015). Hence, they show some potential in being used as diagnostic tools even with the younger age group and before evidence of syntactic knowledge is evident in the child, as is often the case with children experiencing language development delays or language impairment. This is important as early detection of potential language delays

or language impairment could guarantee that the child will receive the much needed attention.

Secondly, the results suggest that the phonological STM plays an important role in the repetition task, but familiarity with the language influences the performance in the NWR tasks. For typically developing children, the NWR task used in this study which was limited to only two to four syllables may not be sensitive enough to discriminate the development of five and six year old children. An in-depth analysis of the errors would be needed to determine the reason(s) why these children were all performing poorly with the four-syllable nonwords. Such an analysis would help further development of the NWR tasks. It is also important to evaluate the consistency of children's performance to further determine its utility as a diagnostic tool for Mandarin-English bilingual children in Malaysia.

The current study was conducted on a very small group of children. To test the efficacy of the NWR task further, a larger sample of children should be involved in future studies. As correctly pointed out by the anonymous reviewers, the children from this study were recruited from only one kindergarten. Coupled with the fact that the sample size is small, the results may not be generalizable to the population of Mandarin-English simultaneous bilingual children in Malaysia. Nonetheless, the current study has shown that the NWR task is an easy task to conduct, and it has potential to be developed as a diagnostic tool to obtain information about bilingual children's language development level without directly tapping the use of specific and explicit grammatical knowledge of children growing up in a complex linguistic environment in Malaysia. This study has only taken the first step towards this goal. More study involving a bigger group of children (typically developing children and children with language development delay as well as language impairment) are needed before any decision can be made about the suitability of this task for diagnosis of children with language impairment.

ACKNOWLEDGEMENT

We would like to thank Dr. Wong Bee Eng and two anonymous reviewers for constructive comments made on earlier versions of this manuscript. We would also like to thank the children, their parents and the teachers from the kindergarten that made this study possible.

REFERENCES

Archibald, L. M. D. (2008). The promise of nonword repetition as a clinical tool. *Canadian Journal of Speech-Language Pathology and Audiology*, *32*(1), 21-28.

Archibald, L. M. D., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders*, *41*(6), 675-693. https://doi.org/10.1080/13682820500442602

A Razak, R., Jin, L., Lim, H. W., & A Aziz, M. A. (2016). Profiling Malay children's syntactic development: A Malay LARSP. In P. Fletcher, M. J. Ball & Crystal, D. (Eds.), *Profiling grammar: More languages of LARSP* (pp. 135-171). Bristol, USA: Multilingual Matters.

A Razak, R., Lim, H. W., A. Aziz, M. A., & Jin, L. (2014, March). *Constructing My-LARSP based on the Malay Developmental Syntactic Profile.* Paper presented at 2014 LARSP Across Languages Workshop, Centre for Literacy & Multilingualism, University of Reading, England.

A Razak, R., Neelagandan, A. I., Md Yusuf, N., Lim, H. W., Ahmad, K., & Madison, C. (2018). The validation of the Malay Preschool Language Assessment Tool (MPLAT): The screening and diagnostic versions. *Malaysian Journal of Public Health Medicine, Special volume*(1), 191-115.

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189-208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: Recent advances in research and theory* (Vol. 8; pp. 47-90). New York, USA: Academic Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*, 575-589. https://doi.org/10.1016/S0022-5371(75)80045-4

Bishop, D. V. M. (2004). Specific language impairment: Diagnostic dilemmas. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp. 309-326). Mahwah, USA: Lawrence Erlbaum Associates.

Bortolini, U., Arfé, B., Caselli, C. M., Degasperi, L., Deevy, P., & Leonard, L. B. (2010). Clinical markers for specific language impairment in Italian: The contribution of clitics and non-word repetition. *International Journal of Language & Communication Disorders/Royal College of Speech & Language Therapists*, *41*(6), 695-712. https://doi.org/10.1080/13682820600570831

Chi, P. S. (2007). Phonological short-term memory in children with specific language impairment. *Bulletin of Special Education*, *32*(4), 19-45. Retrieved May 3, 2013, from http://bse.spe.ntnu.edu.tw/upload/journal/prog/994_3238_26CU_10RY518.pdf

Chiat, S., & Roy, P. (2007). The preschool repetition test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language, and Hearing Research*, *50*, 429-443. doi: 10.1044/1092-4388(2007/030)

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairment (SLI). *International Journal of Language & Communication Disorders*, *43*(1), 1-40. https://doi.org/10.1044/1092-4388(2007/030)

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(6), 741-748. https://doi.org/10.1111/1469-7610.00770

De Bree, E., Rispens, J., & Gerrits, E. (2007). Non-word repetition in Dutch children with (a risk of) dyslexia and SLI. *Clinical Linguistics & Phonetics*, *21*(11-12), 935-944. https://doi.org/10.1080/02699200701576892

De V. Hage, S. R., & Grivol, M. A. (2009). Reference values of nonword repetition test for Brazilian Portuguese-speaking children. *Journal of Applied Oral Sciences*, *17*, 63-68. https://doi.org/10.1590/S1678-77572009000700011

Dispaldro, M., Deevy, P., Altoé, G., Benelli, B., & Leonard, L. B. (2011). A cross-linguistic study of real-word and nonword repetition as predictors of grammatical competence in children with typical language development. *International Journal of Language & Communication Disorders*,

46(5), 564-578. https://doi.org/10.1111/j.1460-6984.2011.00008.x

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research (JSLHR)*, *41*(5), 1136-1146. https://doi.org/10.1044/jslhr.4105.1136

Duanmu, S. (2007). *The phonology of Standard Chinese*. Oxford, England: Oxford University Press.

Ellis-Weismer, S., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language and Hearing Research*, *43*(4), 865-878. https://doi.org/10.1044/jslhr.4304.865

French, L. M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, *29*(3), 463-487. https://doi.org/10.1017/S0142716408080211

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(04), 513-543. https://doi.org/10.1017/S0142716406060383

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200-213. https://doi.org/10.1016/0749-596X(89)90044-2

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*(3), 336-360. https://doi.org/10.1016/0749-596X(90)90004-J

Gathercole, S. E., & Baddeley, A. D. (1996). *The children's test of nonword repetition.* London, England: Psychological Corporation.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, *13*, 65-77. https://doi.org/10.1002/(SICI)1099-0720(199902)13:1%3C65::AID-ACP548%3E3.0.CO;2-O

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Pscholinguistics*, *12*, 349-367.

Girbau, D., & Schwartz, R. G. (2007). Non-word repetition in Spanish-speaking children with specific language impairment (SLI). *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, *42*(1), 59-75. https://doi.org/10.1080/13682820600783210

Graf Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research (JSLHR)*, *50*(1), 177-95. https://doi.org/10.1044/1092-4388(2007/015)

Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*, *24*(02). https://doi.org/10.1017/S0142716403000158

Hoff, E., Core, C., & Bridges, K. (2008). Nonword repetition assesses phonological memory and is related to vocabulary development in 20- to 24-month-olds. *Journal of Child Language*, *35*, 1-14. doi: 10.1017/s0305000908008751

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology*, *38*, 241-253. https://doi.org/10.1016/0022-0965(84)90124-3

Ibrahim, H. M., Lim, H.-W., Ahmad Rusli, Y., & Lim, C.-T. (2019). Speech stimuli and nasalance scores for the assessment of resonance in Mandarin speaking Malaysian children. *Clinical Linguistics & Phonetics, 34*(6), 554-565. https://doi.org/10.1080/02699206.2019.1668480

Kosaka, M. (2009). Nonword repetition tasks in Japanese as clinical markers for discrimination between specific language impairment and typically developing children. *Journal of Medical Welfare*, *14*(2), 57-66.

Leonard, L. B. (2000). *Children with specific language impairment*. Cambridge, USA: MIT Press.

Lim, H. W. (2018). Multilingual English-Mandarin-Malay phonological error patterns: An initial cross-sectional study of 2 to 4 years old Malaysian Chinese children, *Clinical Linguistics & Phonetics, 32*(10), 889-912. https://doi.org/10.1080/02699206.2018.1459852

Lim, H. W., Wells, B., & Howard, S. (2015). Rate of multilingual phonological acquisition: Evidence from a cross-sectional study of English-Mandarin-Malay. *Clinical Linguistics & Phonetics, 29*(11), 793-811. https://doi.org/10.3109/02699206.2015.1048379

Masoura, E. V., & Gathercole, S. E. (1999). Phonological short-term memory and foreign language learning. *Internation Journal of Psychology*, *34*(5), 383-388. https://doi.org/10.1080/002075999399738

Masoura, E. V., & Gathercole, S. E. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory*, *13*, 422-429. https://doi.org/10.1080/09658210344000323

Messer, M. H., Leseman, P. P. M., Boom, J., & Mayo, A. Y. (2010). Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology*, *105*, 306-323. https://doi.org/10.1016/j.jecp.2009.12.006

Rispens, J., & Parigger, E. (2010). Non-word repetition in Dutch-speaking children with specific language impairment with and without reading problems. *British Journal of Developmental Psychology*, *28*(1), 177-188. https://doi.org/10.1348/026151009X482633

Stokes, S. F., Wong, A. M.-Y., Fletcher, P., & Leonard, L. B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research (JSLHR)*, *49*(2), 219-36. https://doi.org/10.1044/1092-4388(2006/019)

Summers, C., Bohman, T. M., Gillam, R. B., Peña, E. D., & Bedore, L. M. (2010). Bilingual performance on nonword repetition in Spanish and English. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, *45*(4), 480-93. https://doi.org/10.3109/13682820903198058

Thordardottir, E. (2017). Are background variables good predictors of need for L2 assistance in school? Effects of age, L1, amount, and timing of exposure on Icelandic language and nonword repetition scores. *International Journal of Bilingual Education and Bilingualism*, *23*(4), 400-422. https://doi.org/10.1080/13670050.2017.1358695

Thordardottir, E., & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders*, *46*(1), 1-16. https://doi.org/10.1016/j.jcomdis.2012.08.002

Thorn, A. S. C., & Gathercole, S. E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *The Quarterly Journal of Experimental Psychology*

A: Human Experimental Psychology, 52A(2), 303-324.

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, *52*, 193-204. https://doi.org/10.1016/j.jml.2004.10.003

Windsor, J., Kohnert, K., Lobitz, K. F., & Pham, G. T. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association*, *19*(4), 298-310. https://doi.org/10.1044/1058-0360(2010/09-0064)

Woon, C. P. (2015). *Use of nonword and sentence repetition tasks with Mandarin-English bilingual children in Malaysia* (Unpublished master's dissertation), Universiti Putra Malaysia, Malaysia.

Woon, C. P., Yap, N. T., Lim H. W., & Wong, B. E. (2014). Measuring grammatical development in bilingual Mandarin English speaking children with a sentence repetition task. *Journal of Education and Learning*, *3*(3), 144-157. https://doi.org/10.5539/jel.v3n3p144

Yap, N. T., A Razak, R. Haman, E., Łuniewska, M., & Treffers-Daller, J. (2017). Construction of the Malay cross-linguistic lexical task: A preliminary report. *Language Studies Working Paper, 8,* 47-61.