Preliminary Findings on a Newly Developed Malaysian Multilingual Sentence Repetition Task for Multilingual Chinese Children in Malaysia

Cho KS\textsuperscript{a}, Rogayah AR\textsuperscript{a,b,*}, Norsofiah AB\textsuperscript{c}

\textsuperscript{a}Speech Science Program, Centre of Rehabilitation and Special Needs, Studies, Faculty of Health Science, Universiti Kebangsaan Malaysia
\textsuperscript{b}Dept. of Postgraduate Studies, Faculty of Education, Language & Psychology, SEGi University, Malaysia
\textsuperscript{c}School of Humanities, Universiti Sains Malaysia

ABSTRACT

INTRODUCTION: Even though multilingualism is the basic fabric of linguistic diversity of mainstream communities in Malaysia, there is generally a lack of assessment tool to assess the language abilities of multilingual children at risk of language impairment. This study is aimed at developing a Malaysian Multilingual Sentence Repetition (MMSR) task which can be used to assess the morphosyntactic abilities of multilingual Chinese children speaking three languages i.e., English-Mandarin-Malay. MATERIALS AND METHOD: 10 typically developing Chinese children within the ages of 4 years until 6 years and 11 months old were recruited. Each child was presented with sentences auditorily using a laptop with a microphone headset. The child was then asked to repeat verbatim the sentences they just heard. RESULTS: Findings showed that the subjects performed equally well in English and Mandarin but lower in Malay ($p < 0.01$). Correlation analyses revealed that the older children performed better for English ($p < 0.05$) and Mandarin ($p < 0.01$) sentence repetitions (SRs). Performance on SR tasks decreased with the increase in sentence length and complexity. CONCLUSION: The newly developed multilingual SR task proved reliable and valid as a tool and has the potential to be developed into a standardized test to assess multilingual Malaysian children at risk of language impairment.

INTRODUCTION

Many Malaysians are at least bilingual if not multilingual. Monolinguals are rare.\textsuperscript{1} Malaysia, a country with a population of over 32 million people, is a diverse country with a multi-ethnic and multilingual population: with three major ethnic groups i.e., Malays, Chinese, and Indians and a plethora of minority groups.\textsuperscript{2-4} It is reported that there are over 130 languages spoken in Malaysia. Malaysians generally speak the national and official language i.e., the Malay language or Bahasa Malaysia (BM), while English is used as a second language for science, trade, and communication sectors.\textsuperscript{2} Vernacular schools use pupils’ own language (Chinese, Tamil) as the medium of instruction.\textsuperscript{2}

The challenges in assessing and diagnosing language disorders in children from multilingual backgrounds have been well documented. Multilingualism affects different populations in different countries: in the Southeast Asia region, it affects the mainstream society whereas in Western countries, it affects the marginalized and migrant communities. There is a dearth of local assessment tools particularly multilingual assessment tools. Most assessment tools in the market are imported tests which are normed based on monolingual children in countries such as the US, UK, Hong Kong, etc.\textsuperscript{5-8} This presents linguistic and cultural biases that hinder adequate assessment of language difficulties in multilingual children. This is evidenced by the risks of both over- and under-identification of language disorders in multilingual children.\textsuperscript{9,10} Many children who might be at risk of developmental language disorder (DLD) in Malaysia were left undetected and untreated due to the unavailability of tests to diagnose DLD.\textsuperscript{5} Currently, there are no
standardised language assessments for multilingual Chinese preschool age children that could test different language abilities in a comparable way.

Sentence repetition (SR) task is one of the tools used to measure typical and atypical language development in children. The task consists of comprehension and production processes involving phonological, morphosyntactic, semantic representation as well as the capacity to store and retrieve linguistic material from memory. SR has been extensively described as a useful tool in identifying language processing abilities and weaknesses. SRs are widely used clinically in identifying children with language difficulties, in particular children with developmental language disorder (DLD).

In this article, we describe the development of a Malaysian Multilingual Sentence Repetition task (MMSR) which consists of three SR tasks: Mandarin SR, English SR and Malay SR. We present preliminary findings of the MMSR including its reliability and validity.

**MATERIALS AND METHODS**

**Study Population**

This cross-sectional study involved 10 typically developing preschool Chinese children recruited from kindergartens in the city of Kuala Lumpur. The inclusion criteria were: i) Malaysian-Chinese ethnicity, ii) within the preschool age of 4 years to 6 years 11 months, iii) exposed to these three languages viz. Malay, English, Mandarin/Chinese dialects, iv) dominant in English or Mandarin as reported by parents, and v) has robust exposure to Malay for at least 6 months. Children who were reported by parents or teachers as suspected of having language delay or any known developmental disorders were not included in the study.

**Study Procedure**

**Test Administration**

Parents were provided with the information sheet and consent form to fill in. Parents also completed the Parent Bilingual Questionnaire (PaBiQ). The study was conducted based on the methodology by Marinis and Armon-Lotem. Children were tested individually in a quiet room at their respective kindergartens. The stimuli were presented using a laptop with a microphone headset. The child was asked to listen to the pre-recorded sentences and repeat the sentences verbatim. All sessions were audio-recorded.

**Screening Instruments**

The baseline language abilities of the participants were screened based on: i) the child’s year end exam school results (Malay, English, Mandarin), ii) a measure of vocabulary in 3 languages based on the Bilingual Language Assessment Battery (BLAB), and iii) a multilingualism survey Questionnaire for Parents of Bilingual Children (PaBiQ). The BLAB assesses the receptive and expressive vocabulary skills of the participants in three languages i.e., Mandarin, English, and Malay. The BLAB Expressive Vocabulary has 60 pictures.

**Research Instrument**

The Malaysian Multilingual Sentence Repetition (MMSR) task is a collation of the Malay SR and the bilingual English-Malay SR. Adaptation of the Mandarin SR tasks was based on these two English-Malay SRs. Sentences were balanced in terms of structure (simple, compound) and sentence types (active, passives, object and subject relative clauses, cleft). Sentences were also arranged according to the levels of complexity; i) Level 1: simple sentence e.g., Subject-Verb-Object, (ii) Level 2: sentences with complements, and iii) Level 3: subordinate sentences, compound sentences, and subject & object relative clauses.

The newly constructed Mandarin SR was validated by a panel of experts. A pilot study of the Malaysian Multilingual SR tasks was carried out on two children.

**Scoring**

The scoring sheet was divided into categorized schemes: i) accuracy, ii) number of content words (content word score), iii) function words (function word score), iv) grammaticality, and v) sentence structure score.
Statistical analysis

Non-parametric tests were used since samples did not assume normality. Face validity was measured with Cronbach’s alpha test. Test-retest and inter-rater reliability were obtained as reliability measures for the Malaysian Multilingual Sentence Repetition (MMSR) task. Examination of the effects of demographic and language variables on children’s performance used Spearman’s rho analyses. Differences in language performance across the SR tasks were examined using Friedman’s ANOVA with post-hoc using Wilcoxon signed rank tests.

RESULTS

Table I (from PaBIQ and BLAB) presents details such as mean age and language profiles of the participants: the average age was 63.5 months (5 years 3 months old). 5 children were between 4 years 0 month and 4 years 11 months old age range, 4 children in the 5 years 0 month and 5 years 11 months age range and 1 child was in the 6 years 0 month and 6 years 11 month age range. Age of onset (AoO) defined as the age at which the child was first exposed to a language, showed that on average, the earliest AoO was English, followed by Mandarin and much later by Malay. For Total Use (TU): English was the most used language, followed by Mandarin and the least used was Malay. A score of 0 means the language is not used at all while a score of 1 signify the maximum usage of a language in daily interactions. Under vocabulary, the numbers signify the scores of the BLAB vocabulary test under each language, a low score signifies more incorrect response and a high score means more correct responses.

Table II shows the performance in terms of mean score and standard deviations together with the Friedman’s ANOVA results. The findings were significant for each score in the multilingual SR task with large effect sizes (Kendall’s W > 0.6).

The post-hoc comparisons reported in Table III indicated that the scores between English and Malay, and the scores between Mandarin and Malay differed significantly on each measure with large size effects (r > 0.5) except vocabulary scores between English and Mandarin. In overall accuracy, they performed similarly. For grammaticality and sentence structures scheme, the participants had equal scores between English and Mandarin but significantly lower scores for Malay.

Age serves as a strong and moderate factor for English and Mandarin SR scores, respectively. Vocabulary scores had positive correlations on English SR scores only. There were no positive correlations of age, AoO, and TU on Malay scores.

Figure 1 shows the differences between accuracy and grammaticality scores. Three sets of Wilcoxon Signed Ranks test revealed significant differences between the English accuracy (M=2.55) and English grammaticality (M=2.50) scores (T=0.00, Z = -2.680, p = 0.007) with a large effect size r=0.599, whereas no significance was reached between Mandarin accuracy (M=2.40) and grammaticality (M=2.45) scores (T=7.00, Z = -1.844, p =0.065), and between Malay accuracy (M=1.05) and grammaticality (M=1.05) (T=0.00, Z=-2.375, p= 0.018).

Table I Age and language profile of participants based on PaBIQ & BLAB screening instruments

<table>
<thead>
<tr>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of subjects (months)</td>
<td>54.00</td>
<td>83.00</td>
</tr>
<tr>
<td>AoO: Age when child is first exposed to a language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>0.00</td>
<td>36.00</td>
</tr>
<tr>
<td>Mandarin</td>
<td>0.00</td>
<td>42.00</td>
</tr>
<tr>
<td>Malay</td>
<td>6.00</td>
<td>48.00</td>
</tr>
<tr>
<td>Total use: Most used language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>0.25</td>
<td>1.00</td>
</tr>
<tr>
<td>Mandarin</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Malay</td>
<td>0.00</td>
<td>0.44</td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>56.00</td>
<td>89.00</td>
</tr>
<tr>
<td>Mandarin</td>
<td>13.00</td>
<td>74.00</td>
</tr>
<tr>
<td>Malay</td>
<td>15.00</td>
<td>35.00</td>
</tr>
</tbody>
</table>

Figure 1. Median accuracy scores against grammaticality scores across the 3 languages
Table II Performance on SR scores (mean and standard deviation, SD) with statistical comparisons between domains in the three languages.

<table>
<thead>
<tr>
<th>Overall accuracy</th>
<th>English (n= 10)</th>
<th>Mandarin (n= 10)</th>
<th>Malay (n= 10)</th>
<th>X², p, W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total score</td>
<td>11.50</td>
<td>11.60</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>(Maximum= 24)</td>
<td>(7.11)</td>
<td>(5.21)</td>
<td>(0.71)</td>
<td></td>
</tr>
<tr>
<td>Mean Rank</td>
<td>2.55</td>
<td>2.40</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>Grammaticality</td>
<td>15.50</td>
<td>15.40</td>
<td>2.80</td>
<td></td>
</tr>
<tr>
<td>(Maximum=24)</td>
<td>(6.62)</td>
<td>(6.79)</td>
<td>(3.29)</td>
<td></td>
</tr>
<tr>
<td>Mean Rank</td>
<td>2.50</td>
<td>2.40</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>Sentence Structure</td>
<td>16.90</td>
<td>16.40</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>(Maximum=24)</td>
<td>(5.55)</td>
<td>(6.42)</td>
<td>(0.94)</td>
<td></td>
</tr>
<tr>
<td>Mean Rank</td>
<td>2.55</td>
<td>2.40</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>Content words</td>
<td>77.35</td>
<td>71.63</td>
<td>23.95</td>
<td></td>
</tr>
<tr>
<td>(Maximum=100%)</td>
<td>(21.49)</td>
<td>(26.04)</td>
<td>(16.35)</td>
<td></td>
</tr>
<tr>
<td>Mean Rank</td>
<td>2.70</td>
<td>2.30</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Function words</td>
<td>27.00</td>
<td>27.50</td>
<td>12.74</td>
<td></td>
</tr>
<tr>
<td>(Maximum=100%)</td>
<td>(24.07)</td>
<td>(27.66)</td>
<td>(10.47)</td>
<td></td>
</tr>
<tr>
<td>Mean Rank</td>
<td>2.50</td>
<td>2.50</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>77.30</td>
<td>52.30</td>
<td>20.60</td>
<td></td>
</tr>
<tr>
<td>(Maximum=90)</td>
<td>(11.79)</td>
<td>(16.03)</td>
<td>(5.68)</td>
<td></td>
</tr>
<tr>
<td>Mean Rank</td>
<td>3.00</td>
<td>1.90</td>
<td>1.10</td>
<td></td>
</tr>
</tbody>
</table>

Note: Friedman’s ANOVA were run to obtain differences. Significant effects are presented in boldface.

Table III Posthoc comparisons of SR domains in the three languages.

<table>
<thead>
<tr>
<th>Scores</th>
<th>English – Mandarin</th>
<th>English – Malay</th>
<th>Mandarin – Malay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>Z</td>
<td>p</td>
</tr>
<tr>
<td>Accuracy</td>
<td>19.50</td>
<td>0.359</td>
<td>0.720</td>
</tr>
<tr>
<td>Grammaticality</td>
<td>25.50</td>
<td>0.205</td>
<td>0.838</td>
</tr>
<tr>
<td>Sentence structures</td>
<td>17.00</td>
<td>-0.637</td>
<td>0.911</td>
</tr>
<tr>
<td>Content words</td>
<td>16.00</td>
<td>-1.722</td>
<td>0.241</td>
</tr>
<tr>
<td>Function words</td>
<td>27.00</td>
<td>-0.051</td>
<td>0.959</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.00</td>
<td>-2.812</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Table III: Significant contrasts indicated in boldface, with effect size, r. The p values indicate the probability of two-tailed tests with the Bonferroni correction of p <0.017 for multiple comparisons.

Figure 2 presents the Friedman’s ANOVA for the differences in repeating sentence structures at different complexity levels which were significant in English (χ²(2) = 7.600 p = 0.022, W = 0.380) and Mandarin (χ²(2) = 11.655, p = 0.003, W = 0.583) but not in Malay (χ²(2) = 3.600, p = 0.165, W = 0.180). Three sets of Wilcoxon Signed Ranks test revealed significant differences between the English level 1 (M= 2.65) & level 3 (M=1.55) scores (T= 0.00, Z = -2.524, p = 0.012) with a large effect size r = 0.564, whereas no significance was reached between levels 1 & 2 (M= 1.80) (T= 4.00, Z = -2.192, p = 0.028) and between levels 2 & 3 (T=
13.00, $Z = -1.126, p=0.260$). These results suggested that the children were significantly more able to repeat sentence structures in English and Mandarin at level 1 and 2 but not at level 3.

Validity and reliability of the Multilingual SR

Face validity of the adapted Mandarin SR showed a Cronbach alpha value of 0.720 or 72% agreement. The Multilingual SR task had good (75.2%) to excellent (100%) for intra and inter-reliability.

Table IV presents correlations between language variables: age, AoO, TU & vocabulary scores on SR scores.

<table>
<thead>
<tr>
<th>Demographic and language variables</th>
<th>English SR</th>
<th>Mandarin SR</th>
<th>Malay SR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.848**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AoO</td>
<td>-0.016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Use</td>
<td>0.421</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.692*</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mandarin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.644*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AoO</td>
<td>-0.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Use</td>
<td>0.134</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.420</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Malay</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.595</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AoO</td>
<td>-0.558</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Use</td>
<td>0.330</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.499</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).**

*Correlation is significant at the 0.05 level (2-tailed).**

**DISCUSSION**

In general, multilingual Chinese children demonstrated equivalent performance in Mandarin and English SRs, but significantly lower in Malay SR.

Chronological age was significantly correlated only with English and Mandarin SR scores. The English and Mandarin scores increased with the increase in age as in past studies. Malay scores, however, were not correlated with the increase in age due to the low total use at home. Age effects on sentence repetition tasks are not consistently reported in other studies. In the Antonijevic-Elliot et al study, they reported no effects of age on bilingual English-Irish children. Only English vocabulary scores had positive correlations with the English SR and this is supported by previous reports in that children with higher vocabulary repertoires have better language skills. Contrary to these studies, vocabulary scores did not correlate to SR scores in Mandarin and Malay due to its low scores. Total use and AoO scores from the PaBiQ parent survey had no positive effects on SR scores in contrast to previous studies. This could be due to the small sample size and low variability of the sample. In general, grammaticality scores between Mandarin and English were not significantly different indicating equal development of the two grammar systems. However, grammaticality scores were significantly higher than accuracy scores for English. This might be due to the additional working memory and sentence processing involved in SR tasks. 2 reasons might be at work: either the grammatical knowledge in English has not achieved maturity and hindered the participants from modifying their answers to suit the targeted English structures or the inadequate or overloaded cognitive capacity. The former seems to be a more plausible explanation. This is because accuracy scores were on par with the grammaticality scores for Mandarin.

With regards to language dominance, participants scored better in their dominant language as measured by total use at home. Across the board, the children scored equally in English and Mandarin SRs. This might be contributed by similar total use of each language at home and the age at which they were first exposed to both languages compared to Malay. The children in this study performed best in the languages they used both at home and school settings similar to past studies.

The significantly better performance of our participants in English and Mandarin SRs contrasted with findings of Woon et al. Woon found that the Malaysian Chinese children in her study had significantly better
SR scores in Mandarin than in English. One possible difference was the sentences used in Woon’s study had different sentence structures tested between Mandarin and English SRs and did not include language independent structures. In contrast, our sentences were comparable in structures and had language independent and language specific structures in all three languages. This is in line with bilingual SR studies tasks which recommended having comparable versions in both/all languages involved when assessing multilingual children.9

In terms of complexity, participants were most successful in repeating sentence structures at complexity levels 1, 2 and least successful at level 3 in English and Mandarin. This was not surprising as previous research have indicated certain structures are acquired later by monolingual and multilingual Chinese children such as the relative clause structure.27 Participants were able to repeat only a few sentences in Malay at level 1. The lower scores were probably due to limited vocabulary and limited exposure to Malay. Generally, performance decreased with the increase in complexity and length as reported by previous studies.23

From the sociolinguistics perspective, the Chinese communities prefer to use English and Mandarin in their home and school environment due to their Chinese heritage and cultural identity.3,33 Their children have limited exposure to Malay: only during school hours or if they have a live-in domestic helper from Indonesia. In addition, language choices in cross ethnic groups interactions showed that Malaysians tended to prefer using English even where the majority of speakers are non-Chinese and opting to speak English rather than Malay.34 Malay adults seemed to prefer to use English especially in urban areas. The quality and quantity of the Malay language input as Language 3 (L3) for Chinese children in private kindergartens received was not comparable to the kind of exposure that children who attend government kindergartens or national type schools where Malay is used as the medium of instruction. These factors can be used to explain the low input of Malay that the Chinese children received in comparison to the high input of Mandarin and English.

CONCLUSION

The current study demonstrated that the Malaysian Multilingual SR (MMSR) task is reliable and valid in assessing multilingual abilities of Malaysian-Chinese children. The comparable structures tested in the language grouping of English-Mandarin-Malay proved viable to test the language abilities of Chinese Malaysian pre-schoolers. The present findings suggest that the Malaysian Multilingual SR is developmentally sensitive. Results from the administration of the SR tasks could assist speech-language therapists in developing remedial targets (for example specific words and sentence structures) in therapy and can be integrated with findings from the other types of assessments done. Future research could focus on developing the MMSR tool into a standardized Multilingual SR to be used on assessing all preschool children in Malaysia.

ACKNOWLEDGEMENTS

Our appreciation goes to the parents, preschool children participants and kindergarten principals who participated in this study. A note of thanks to Dr Susan Rickard Liaw, Director of Speech & Language Pathology Program, Faculty of Medicine, National University of Singapore by providing us access to the BLAB test.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

REFERENCES


