

Kertas Asli/Original Article

**Mandarin Speech Perception Tests for Malaysian Chinese Children
(Ujian Persepsi Pertuturan Mandarin untuk Kanak-kanak Berbangsa Cina di Malaysia)**

CILA UMAT, CHONG SHEUE LIH & SITI ZAMRATOL-MAI SARAH MUKARI

ABSTRACT

The study aimed to develop Mandarin speech perception tests for Malaysian Chinese children. Two tests were developed which were the Syllabic Pattern Perception Test (SPPT) and the Tone Perception Test (TPT). Both tests were designed for children aged between three to six years-old. A total of 80 normal hearing children aged 3 years to 6 years and 11 months from several kindergartens around Kuala Lumpur and Ipoh participated in this study. Results showed that the mean scores for the SPPT did not differ significantly across age ($p > 0.05$), possibly due to a ceiling effect. Although the SPPT showed relatively low test-retest and inter-rater Spearman correlation coefficients ($r = 0.68$ for correct syllable pattern response and $r = 0.38$ for correct item response), the scores were highly repeatable in both test-retest and inter-rater conditions. For the TPT, the mean scores of five and six year old children were significantly higher than three year old subjects while the mean score of four year old subjects was significantly poorer than the six years old. The test-retest and inter-rater reliability were high and the coefficients approximated 0.80 ($p < 0.05$). For both tests, gender and dialects or languages used as the second language did not have significant effect on the test scores. Content validity evidence was also collected from a group of panelists. In conclusion, SPPT and TPT are reliable and valid Mandarin speech tests to be used on Malaysian Chinese children who speak Mandarin.

Keywords: Malaysian Chinese children, Mandarin speech tests, tone perception, syllabic pattern perception

ABSTRAK

Kajian ini bertujuan untuk membentuk ujian pertuturan Mandarin untuk kanak-kanak berbangsa Cina di Malaysia. Dua ujian telah dibentuk dalam kajian ini iaitu Ujian Persepsi Corak Sukukata (SPPT) dan Ujian Persepsi Nada (TPT). Kedua-dua ujian dibentuk untuk menguji persepsi pertuturan kanak-kanak berusia di antara 3 hingga 6 tahun. Seramai 80 kanak-kanak berpendengaran normal berumur antara 3 tahun hingga 6 tahun 11 bulan dari beberapa buah tadika sekitar Kuala Lumpur dan Ipoh menyertai kajian ini. Hasil kajian menunjukkan bagi ujian SPPT, tidak terdapat perbezaan skor purata yang signifikan di antara kumpulan umur ($p > 0.05$), kemungkinan disebabkan kesan siling. Kebolehulangan skor semasa uji-uji semula dan antara penguji adalah sangat tinggi bagi SPPT walaupun koefisien kedua-duanya rendah secara relatif ($r = 0.68$ untuk respons corak suku kata betul dan $r = 0.38$ untuk respons item betul). Bagi TPT, min skor di kalangan kumpulan umur lima dan enam tahun adalah lebih tinggi dari kumpulan umur tiga tahun sementara min skor subjek empat tahun lebih rendah secara signifikan berbanding min skor kumpulan enam tahun. Kebolehpercayaan uji-uji semula dan antara penguji adalah tinggi di mana koefisien kedua-duanya menghampiri 0.80 ($p < 0.05$). Bagi kedua-dua ujian, jantina dan dialek atau bahasa kedua yang digunakan tidak mempunyai kesan yang signifikan terhadap skor yang diperolehi. Bukti kesahan kandungan juga diperolehi dari sekumpulan ahli panel. Kesimpulannya, SPPT dan TPT adalah ujian pertuturan Mandarin yang boleh dipercayai dan sah untuk digunakan pada kanak-kanak Cina berbahasa Mandarin di Malaysia.

Kata kunci: Kanak-kanak Cina Malaysia, ujian pertuturan Mandarin, persepsi tonal, persepsi corak sukukata

INTRODUCTION

Speech perception tests are widely used in audiology to measure the auditory perceptual capability of the hearing-impaired population (Mendel & Danhauer 1996; Derinsu et al. 2007). Measurement of speech perception, together with other procedures, allow audiologists to assess the benefits of hearing aid and/or cochlear implant usage and to determine the needs for cochlear implant in children and adults. Clark et al. (1997) for example, recommended speech

perception tests to be used in determining cochlear implant candidacy in children. Speech perception tests also help in setting goals in aural rehabilitation (Eisenberg et al. 2005).

In Malaysia, several speech tests have been adapted into Malay for the use of local Malay population. These include the Malay Speech Audiometry test which consists of two-syllable word lists (Mukari & Said 1989), the Malay Hearing in Noise Test (MyHINT) (Quar et al. 2008), and the Malay version of Evaluation of Auditory Responses to Speech (EARS) (Mukari & Hamid 2008). However, Malaysia

is a multiracial country, and yet, to date, there is still no speech test in the language of other major races in Malaysia. Carter et al. (2001) stated that children's performance in speech tests will be influenced by the language of the speech tests. Thus, speech tests in Malay may not be suitable to assess the Chinese population in Malaysia as Malay is not their native and dominant language. As of June 2008, it was reported that there were a total of 7,150,000 Chinese in Malaysia (Wikipedia 2009). It is important to develop speech perception tests in Mandarin that are culturally appropriate to the local population of Chinese children since the Chinese commonly makes up 26% of the Malaysian population (Wikipedia 2009). The development of speech perception tests for the Chinese community is challenging due to the multiple dialects spoken by this population. However, since Mandarin is the language of instruction in Chinese schools across Malaysia, speech perception tests in Mandarin are in demand.

For children with limited listening skills, speech testing and training may focus on syllable pattern perception because pattern perception is the easiest test for children. In such an assessment, subjects only need to detect the stress and temporal patterns of the test items (Derinsu et al. 2007). Assessment of tone perception is also important in Mandarin as it is a tonal language, which uses different tones to discriminate meaning of a syllable (Wong et al. 2005). The inability to discriminate between tones will therefore cause difficulties in the understanding of Mandarin. In view of the above considerations, this study involved the development of speech test materials for Malaysian Chinese children - the Syllabic Pattern Perception Test (SPPT) and the Tone Perception Test (TPT). The tests were targeted for children aged between 3 to 6 years old.

MATERIALS AND METHODS

SUBJECTS

Subjects were Malaysian Chinese children who speak Mandarin as their dominant language, recruited from several kindergartens around Kuala Lumpur and Ipoh. All subjects had normal hearing, speech, vision and physical development as reported by their parents and teachers. The number of participants and subjects' age groups were different for each phase of the study.

In the first phase of the study in which targeted children's vocabulary were collected, 20 children aged two years old participated to ensure items used in the tests would be within the vocabulary of the three year olds, the youngest targeted age group. A total of 15 children aged three years old participated in the second phase, which was the pre-test. No subjects were involved in the third phase in which the test items were selected. A pilot study was conducted in the fourth phase which involved 10 children between three to four years of age. For the last phase, which was the field study of the newly developed

tests, 80 children aged between three to six years old participated. There were 20 children in each age group: 3 years to 3 years and 11 months (3-3.11), 4 years to 4 years and 11 months (4-4.11), 5 years to 5 years and 11 months (5-5.11) and 6 years to 6 years and 11 months (6-6.11).

Since the majority of the Chinese populations in Malaysia speak more than one language or dialect, bilingual children were also included in the study. Forty of the subjects involved in the final phase (field study) participated in the test-retest reliability study and another 40 subjects participated in the inter-rater reliability study.

CONSIDERATIONS IN THE DEVELOPMENT OF THE TESTS

Developing speech perception tests for children is challenging due to the limitation of vocabularies and language. It is important to make sure that test items are within the vocabulary that are mastered by the tested children and the response task should be age-appropriate to ensure correct interpretations (Martin 1987; Mendel & Danhauer 1996; Clark et al. 1997). In the present study, to ensure that the test items are age-appropriate, the items were selected from vocabularies that were mastered by 20 two-year-old children in the first phase and the selected items were also pretested on another group of three-year-old children in the second phase. Besides, the response task for both tests are picture-pointing and the format used are closed-set. This was to ensure that the children's performance in the tests was not affected by the limitation of their expressive language.

There were other considerations taken into account in the selection of test items. Since the main purpose of the SPPT was to determine the ability of test takers to discriminate between different syllable patterns, it was important to make sure that the test items selected exist in only one form of the syllable pattern. In Mandarin, every syllable contributes to the meaning of the word. Some of the words in Mandarin can be understood in monosyllabic, disyllabic or polysyllabic pattern. For instance, elephant in Mandarin can be represented by '象 (elephant)', '大象 (big elephant)' or '笨象 (big and dumb elephant)' in Malaysian Mandarin. Although these words differ in syllable pattern, they bring the same meaning especially for young children. They are all 'elephant.' Thus, items selected to be in the final test list were those that can only exist in one form of the syllable pattern. Additionally, Malaysian Chinese speak more than one language and most of them can speak a simple English and Malay languages. Children as young as three years old are exposed to different languages or dialects other than Mandarin. Some parents often teach their children some simple vocabulary in English. Items that are often spoken in English or other languages were discarded from the test list to avoid test takers give wrong responses as a result of unfamiliarity of the test items in Mandarin.

The selection of test items in the TPT was also complicated. The items selected had to cover the six

possible tone combinations of the four tones in Mandarin. Besides, the test items must be in monosyllabic pattern to eliminate segmental cue. In Mandarin, every syllable can stand alone; yet, it always combines with other syllables to form a polysyllabic word. Although this word differs in syllable pattern, it has the same meaning. For example, ‘帽’ and ‘帽子’ have the same meaning, which is ‘hat’ in English. However, between these two forms of word, ‘帽子’ is more common. For instance, in the pretest stage, the subject tended to misidentify ‘帽’ as ‘猫’ which is the minimal pair of ‘帽’ as a result of the unfamiliarity of the test item in its monosyllabic form. Thus, it is important to ensure that the selected items are in a common form. The selection of items in this study had excluded words that are not common in a monosyllabic pattern and every test items was proven not to be confusing to the subjects.

Although a syllable with different tones conveys different meaning, in Malaysia, it is quite common for some Mandarin speakers to produce some of the syllables in a wrong tone. This is especially so done by those who do not receive education in Chinese primary and secondary schools. Although they can speak Mandarin, their usage of tones may not be correct. In addition, code-switching between different dialects and Mandarin is also common among the Malaysian Chinese population; some of the tones used in Mandarin are confused with other tones used in other dialects. Such confusing items were excluded from the final test lists.

TEST PROCEDURE

ITEM POOL GENERATION

In the first phase of the study, the Mandarin First Word Checklist (Chok 2001) was given to 20 parents or caregivers of children aged 2;0 to 2;11 to gather monosyllabic, bisyllabic and trisyllabic words that were understood or within the vocabulary of these children. Twenty four items (eight for each syllable pattern) that were understood by at least 80% of the subjects were selected for the SPPT and 36 items (three minimal pairs for each tone combination) for the TPT.

The selected items were drawn in picture cards and then pretested on 15 three-year-old subjects in the second phase. Out of the 24 and 36 items selected in the first stage for the SPPT and the TPT, respectively, 12 items for each test that could be recognized by at least 80% of the subjects were selected to be test items. The 12 items chosen for the SPPT cover all the important consonants and vowels in Mandarin, based on the Mandarin Word List (Mok 2000).

Next, the selected test items were arranged according to the following criteria to produce the two tests: items or tone pairs with the highest percentage of correct responses were selected to be in the easiest test lists (level A). For the SPPT, there are three levels of difficulty in which Level A is the easiest while Level C is the hardest (it contains all the

12 test items). As for the TPT, there are only two levels of difficulty (Level A and Level B). In level A, there are two minimal pairs that cover all the four Mandarin tones while there are six minimal tone pairs in Level B. The newly developed tests were then trialed on 10 subjects (five 3-year-olds and five 4-year-olds) and further modifications and refinement were done based on the results of the pilot study before the field study commenced.

TESTS COMPOSITION

The objective of the SPPT is to determine the ability of test takers to discriminate different syllable patterns. The test is comprised of monosyllabic, bisyllabic and trisyllabic words.

Level A, which is the easiest level, consists of three items (that is, one for each syllable pattern), Level B has six items (that is, two for each syllable pattern) and Level C, the most challenging level, has 12 items (four for each syllable pattern). In the SPPT, scorings are based on (1) the correct syllabic pattern response; and (2) the correct item response.

The TPT assesses the ability of test takers to discriminate between the four different tones in Mandarin. The test is comprised of minimal pairs of monosyllabic words that differ in tone only. The items in the test cover the six possible tone combinations of the four tones in Mandarin. As mentioned earlier, Level A is comprised of only two minimal pairs that cover the four tones and Level B comprises of six minimal pairs that cover the entire possible tone combinations. Scorings for the TPT are based on (1) the correct syllable response; (2) the correct item response; and (3) the correct minimal pair response.

Both the SPPT and the TPT are closed-set tests and the response can either be picture-pointing and/or verbal responses.

FIELD-TESTING

Both tests were field-tested on 80 subjects. The presentation of the test items were at random based on the randomization table developed for the tests. The stimuli were presented using live voice and in auditory-alone mode. Subjects were requested to point to the corresponding picture card of the test item after each presentation of the stimulus. Even though the formal instruction of the test was for subjects to point to the correct picture card, verbal responses were also accepted. Repetition of stimulus-presentation during the test was not allowed.

A few practice trials were given to ensure that subjects understood the test procedure and the required responses. All the three test levels in the SPPT and two levels in the TPT were tested and the scores of the subjects were recorded. To obtain information on test-retest and inter-rater reliability, 40 subjects were retested by the same tester and another 40 subjects were retested by a different tester two weeks after the field test sessions.

Both testers, a qualified audiologist and a final year Audiology student, were native Mandarin speakers and received Chinese Mandarin education in primary and secondary schools. The testers were briefed on the testing and scoring procedures.

CONTENT VALIDITY

Content validity evidence of the SPPT and the TPT was collected from a group of panelists that consisted of an audiologist, an undergraduate audiology student and five undergraduate speech sciences students. All members of the panel were native Mandarin speakers and received Chinese Mandarin education in primary and secondary schools. The members had experiences in administering speech tests and had basic knowledge on language development. Content validity was performed prior to the field study.

DATA ANALYSIS

A three-way ANOVA was used to analyze the difference in test scores among different age groups, genders and different languages or dialects used. Spearman Correlation was used to analyze the correlation of the scores of the first test and those of the repeated test in the test-retest reliability study as the data was not normally distributed.

RESULTS

SUBJECTS

There were 10 subjects per age group (3 to 6 years olds) participated in the final phase of the study with a mean age of 4.74 years and a standard deviation (SD) of 1.09 years old. The results described below for both the SPPT and the TPT tests were from these group of subjects.

DESCRIPTIVE ANALYSES

Figure 1 and Figure 2 show the mean scores and the SD obtained by subjects in the different age groups for the SPPT and the TPT, respectively.

As can be seen from Figure 1 and Figure 2, the scores in both tests approximated ceiling level for all levels of difficulty (maximum scores for Level A = 12; Level B = 18 and Level C = 24) and all types of scores. The mean percent correct scores for each of the four Mandarin tones across age groups are illustrated in Figure 3.

In order to determine which tone combination was the most difficult for the subjects, the mean percent correct responses of each of the six tone combination were calculated for all age groups and shown in Table I.

The results showed that three and four-year-old subjects had slight difficulties in discriminating two of the tone combinations with scores around 70% or less. Combination of rising (tone 2) versus dipping tone (tone 3)

was the most difficult for the 3year-olds (mean score = $68.8\% \pm 38.8\%$) while combination of level (tone 1) versus dipping tone (tone 3) was the most difficult for the 4 year olds (mean score = $71.3\% \pm 30.7\%$). For five and six-year-old subjects, all of the tone combination were perceived with high accuracy in which the mean percentage correct of the entire tone combinations exceeded 85%.

THE EFFECTS OF AGE, GENDER AND DIFFERENT DIALECTS OR LANGUAGE GROUPS

Three-way ANOVA showed that there was no significant difference between the performance of different age groups [$F(3, 78) = 0.95, p > 0.05$], different gender [$F(1, 78) = 0.25, p > 0.05$], and different dialects used [$F(2, 78) = 0.65, p > 0.05$] for the SPPT. This was possibly due to the ceiling effect as majority of the subjects achieved almost 100% score. However, for the TPT, there was a significant difference between the different age groups [$F(3, 78) = 4.41, p < 0.05$] but not for the different gender [$F(1, 78) = 0.09, p > 0.05$] and different languages or dialects used [$F(2, 78) = 0.04, p > 0.05$]. Post-hoc Dunnett T3 showed that the mean score of three-year-old subjects was significantly poorer than the five- and six-year old subjects while the mean score of four-year-old subjects was significantly poorer than the six-year-olds.

TEST-RETEST AND INTER-RATER RELIABILITY

Pearson correlation was used to determine the relationship between the test scores obtained on the first and the repeated tests. A strong reliability is defined as a minimum of 0.80 correlation coefficient obtained (Agency for Healthcare Research and Quality 2002).

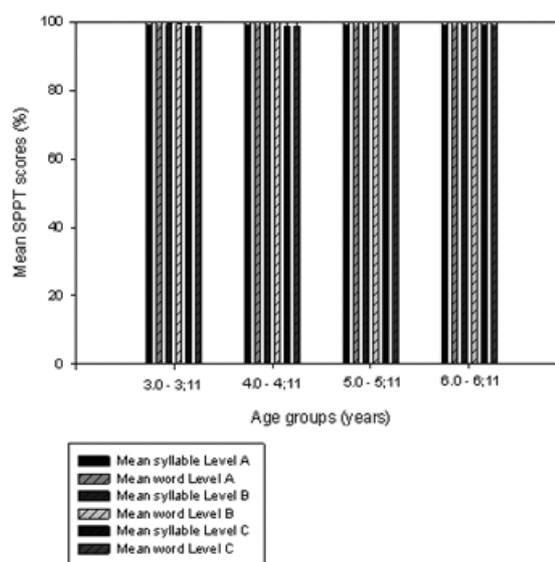


FIGURE 1. Mean scores for correct syllable patterns and correct words for each level of difficulties in the Syllable Pattern Perception Test (SPPT) across age groups

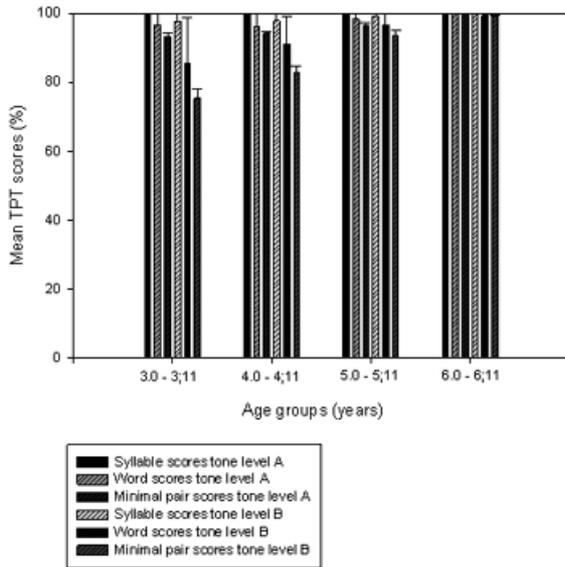


FIGURE 2. Mean scores for correct syllable patterns, correct words and correct tone minimal pairs for each level of difficulties in the Tone Perception Test (TPT) across age groups

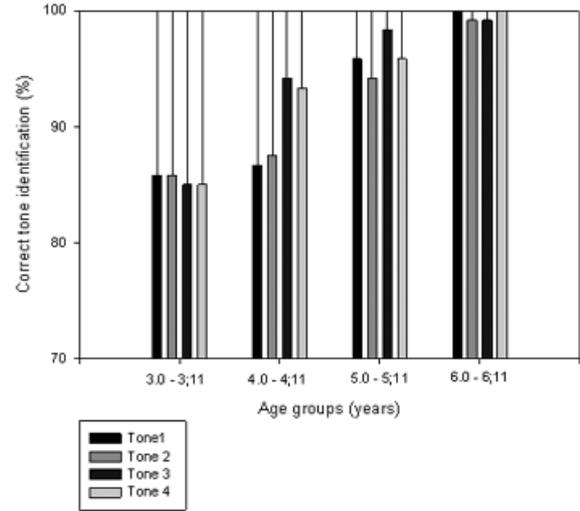


FIGURE 3. Mean scores for correct tone identification (in %) for the four Mandarin tones across age groups

TABLE 1. Mean scores (in %) and the standard deviations (SD) for tone minimal pairs across age groups in the TPT

Tone combinations	3 – 3;11		4 – 4;11		5 – 5;11		6 – 6;11	
	Mean (%)	SD (%)						
1 & 2	81.3	29.1	78.8	33.7	91.3	20.3	100.0	0.0
1 & 3	71.3	38.3	71.3	30.7	96.3	16.8	98.8	5.6
1 & 4	80.0	34.0	95.0	15.4	98.8	5.6	100.0	0.0
2 & 3	68.8	38.8	87.5	30.9	85.0	31.8	97.5	11.2
2 & 4	76.3	32.9	73.8	38.5	90.0	24.9	98.8	5.6
3 & 4	92.5	16.4	97.5	7.7	100.0	0.0	100.0	0.0

For Levels A and B in the SPPT, as all subjects obtained 100% scores, the correlation coefficients for both test-retest and inter-rater reliability could not be established. In Level C, the test-retest and inter-rater correlation coefficients for both the correct syllable pattern response and the correct item response were $r = 0.68$ and $r = 0.38$, respectively with $p < 0.05$. The mean scores obtained in the first and repeated tests for Level C across the age groups differed by 0.12% in the test-retest and no difference in the inter-rater condition.

As for the TPT, for Level A, the correct item response and the correct minimal pair response showed strong and significant test-retest correlation coefficients ($r = 0.82$, $p < 0.05$). The inter-rater correlation coefficient for the correct item response was $r = 0.79$, $p < 0.05$ and the correct minimal pair response was $r = 0.77$, $p < 0.05$. For the correct syllable response, the correlation coefficients for both test-retest and inter-rater reliability could not be calculated as all of the subjects obtained similar scores (100%).

In Level B of the TPT, the test-retest correlation coefficients for the correct syllable response, the correct

item response and the correct minimal pair response were $r = 0.44$ ($p < 0.05$), $r = 0.72$ ($p < 0.05$) and $r = 0.71$ ($p < 0.05$), respectively. In the inter-rater condition, the correlation coefficient for both the correct item response and the correct minimal pair response was $r = 0.84$, $p < 0.05$.

CONTENT VALIDITY

All of the members of the panel (100%) rated the SPPT as a suitable test to assess speech syllable pattern perception of children between three to six years of age. Approximately 87.5% of the members in the panel rated TPT as suitable to examine tone perception of children in the targeted groups. Some panel members suggested that certain items in the TPT were difficult for the three-year-olds. Both of the tests were modified based on the panel's comments.

DISCUSSION

This study reported the development of two closed-set Mandarin speech perception tests for children aged 3 to 6 years old, namely the SPPT and the TPT. The test-retest and inter-rater reliability were also examined.

Results revealed that all subjects performed well in the SPPT. In such a test, subjects need to be able to discriminate between stress and temporal pattern of the items (Derinsu et al. 2007), which is not an issue in children with normal hearing and intact auditory system pathway. Moreover, the test items were selected from the vocabularies that were mastered by children as young as two years old. Hence, the SPPT is suitable for the assessment of children aged between three to six years old.

A previous study reported by Wong et al. (2005) showed that three years old children had no difficulty in discriminating and identifying the four Mandarin tones with the presence of segmental cues, but had difficulties when the segmental cues were eliminated. Hence, the relatively low performance in some of the tone combinations for children in the younger age groups might be due to the fact that some subjects were not able to identify the four Mandarin tones when segmental cues were not present. All items used in the tone pairs in this study are monosyllables. Thus, potential users of these tests should be cautious when interpreting the results from three and four years old children as the present study showed even some of normal hearing children had not yet mastered some of the tone combinations. However, when the items were not in minimal pairs, the children had no difficulties in the identification or discrimination of the items as the vocabularies used were within the speech and language repertoire of these younger children.

The findings that age affected tone perception scores are consistent with several earlier studies related to speech and language (Jerger et al. 1983; Merckham & Hazan 2004; Bradley & Sato 2004; Talarico et al. 2007) in which mean scores increased with age. The age effect was not observed in the SPPT as almost all subjects obtained 100% scores. As for the nil gender effect in both tests, the results supported in part the finding by Norris et al. (1989) in which there were no significant difference between gender on children's speech and language performance. However, Karmiloff & Karmiloff-Smith (2002) reported that girls were faster and better than boys in speech and language development. The present results also showed that there was no significant difference between the mean scores of subjects that used different languages or dialects as their second language, possibly as a result of a ceiling effect. This finding is inconsistent with a previous study which reported that bilingual children who spoke both Mandarin and Cantonese have better tone awareness than those who were monolingual (Chen et al. 2004).

The criterion of test-retest and inter-rater reliability were met if the correlation coefficient between tests were

0.90 and above (McClauley & Swisher 1984). However, some might argue that the 90% criterion for reliability was too high given the complexity of speech and language functioning and disorders. Additionally, the variability in daily performance that arises from different speech and language disorders suggests that 0.90 criterion is fairly high. Thus, the Agency for Healthcare Research and Quality (2002) suggested that correlation coefficient as low as 0.80 was considered as the threshold of acceptability for reliability. Thus, the present study set 0.80 correlation coefficients as the standard for strong reliability. In any study involving test-retest and inter rater reliability, an appropriate length of test-retest interval is important. A short interval between the tests will lead to learning/practice effect, yet, a long interval between the tests might be invalid as a result of maturation effect (Garson 2008). Garson (2008) suggested that a typical interval is several weeks. Hence, in the present study, the test-retest interval selected was two weeks.

The reliability study showed that the test-retest and inter-rater correlation coefficients for level C in SPPT were lower than the set criterion for correlation coefficient (0.80). However, the test was still highly repeatable as 97.5% of the subjects obtained the same score in the repeat test after the two weeks interval and 92.5% of the subjects obtained the same score in the repeat test done by the second tester. In both Level A and level B of the SPPT, all subjects were able to repeat their scores in the first and the repeated test. As for the TPT, the test-retest and inter-rater correlation coefficients approximated the criterion correlation coefficient for reliability (0.80), which were 0.71 and 0.82, respectively. The scores obtained in the first and the repeat tests were very similar, in which the differences ranked between one to three marks. The above findings indicate that both tests are highly reliable to be used at different times and conducted by different testers.

CONCLUSION

Two Mandarin speech perception tests were developed in this study to quantify the ability of Malaysian Chinese children aged between three to six-years-old to perceive syllabic pattern differences and differences in Mandarin tones. As all the items chosen for the tone perception test are monosyllabic, the study showed that some three and four-year-old children had not fully mastered the four tones in Mandarin without segmental cues. Hence, TPT test users should be cautious when interpreting results of three and four-year-old children as even some normal hearing children still had not fully mastered the four tones. There was a significant effect of age in the TPT but not in the SPPT. Gender and dialects or languages used as second language did not affect the test scores of both tests. The tests were also found to be reliable and valid.

ACKNOWLEDGEMENT

The authors thank all children and parents who had given the consent for their children to participate in the study. The study was funded by the e-Sciencefund research grant (06-01-02-SF0449) from the Ministry of Science, Technology and Innovation (MOSTI). The e-Sciencefund project has received approval from the Human Ethics Committee, Universiti Kebangsaan Malaysia to be conducted on human subjects.

REFERENCES

- Agency for Healthcare Research and Quality. 2002. Criteria for determining disability in speech-language disorders. *Evidence Report/technology Assessment* 52: 1530-440X
- Bradley, J.S. & Sato, H. 2004. *Speech Intelligibility Test Results for Grades 1,3 and 6 Children in Real Classrooms*. National Research Council, Canada and Tohoku University, Japan
- Carter, J.A., Lee, J.A., Murira, G.M., Gona, J., Neville, B.G.R. & Newton, C.R.J.C. 2001. Review: Issues in the development of cross-cultural assessments of speech and language for children. *Int. J. Lang. Comm. Dis.* 40:385-401.
- Chen, X., Richard, C., Li, W.L., Hao, M.L., Wu, X.C. & Shu, H. 2004. Phonological awareness of bilingual and monolingual Chinese children. *J Ed Psychol.* 96: 142-151.
- Chok, S.S. 2001. *Perkataan pertama dalam Bahasa Mandarin yang dihasilkan oleh kanak-kanak berbangsa Cina di Malaysia*. Unpublished Bachelor of Speech Sciences thesis, Kuala Lumpur: Universiti Kebangsaan Malaysia.
- Clark, G.M., Cowan, R.S.C. & Dowell, R.C. 1997. *Cochlear Implantation for Infants and Children - Advances*. United States: Singular Publishing Group, Inc.
- Derinsu, U., Gumus, Z.G. & Akdas, F. 2007. Standardization of the speech tests used with cochlear implant patients. *Marmara Medic J.* 20: 6-11.
- Eisenberg L.S., Johnson, K.C. & Martinez, A.S. 2005. *Clinical Assessment of Speech Perception for Infant and Toddlers*. Los Angeles: Children's Auditory Research and Evaluation Centre, House Ear Institute.
- Garson, D. 2008. *Reliability: Reliability Analysis*. Retrieved April 13, 2008 from <http://www2.chass.nsu.edu/garson/pg765/reliab.htm#retest>
- Jerger, S., Jerger, J. & Abrams, S. 1983. Speech audiometry in the young children. *Ear Hear.* 4: 56-66.
- Karmiloff, K. & Karmiloff-Smith, A. 2002. *Pathways to Language*. United States: Harvard University Press.
- Martin M. 1987. *Speech Audiometry*. United Kingdom: Taylor & Francis Ltd.
- McClauley, R.J. & Swisher, L. 1984. Psychometric review of language and articulation tests for preschool children. *J. Sp. Hear. Dis.* 49: 34-42.
- Mendel L.L. & Danhauer J.L. 1996. *Audiologic Evaluation and Management and Speech Perception Assessment*. San Diego: Singular Publishing Group, Inc.
- Merkham, D. & Hazan, V. 2004. The effect of talker and listener-related on intelligibility for a real word, open-set perception test. *J. Sp. Lang. Hear. Res.* 47: 725-737.
- Mok, Y.Y. 2000. *Senarai perkataan Bahasa Mandarin*. Unpublished Bachelor of Audiology thesis. Kuala Lumpur: Universiti Kebangsaan Malaysia.
- Mukari, S.Z.M. & Said, H. 1989. The development of Malay speech audiometry. *Med. J Malaysia* 46(3): 262-268.
- Mukari, S.Z.-M.S. & Hamid, B.A. 2008. *Penilaian Respons Auditori kepada Pertuturan*. Bangi: Penerbit Universiti Kebangsaan Malaysia.
- Norris, D., Cutler, A., Mehler, J. & Segui, J. 1989. *Limits on bilingualism*. *Nature* 340: 229-230.
- Quar T.K., Mukari S.Z., Abdul Wahab, N.A, Abdul Razak R, Omar M. & Maamor N. 2008. The Malay Hearing in Noise Test. *Int. J. Audiology* 47: 379-380.
- Talarico, M., Abdilla, G. Aliferis M., Balazic, I., Giaprakis, I., Stefanakis, T., Foenander, K., Grayden, D.B. & Paolini, A.G. 2007. Effect of age and cognition in speech and noise perception abilities. *Audio Neurotol.* 12: 12-19.
- Wikipedia. 2009. *Malaysian Chinese*. Retrieved Feb 28, 2009 from http://en.wikipedia.org/w/index.php?title=Malaysian_Chinese.
- Wong, P.S., Schwartz, R.G. & Jenkins, J.J. 2005. Perception and production of lexical tones by 3 Year-Old, Mandarin speaking children. *J. Sp. Lang. Hear. Res.* 48: 1065.

Cila Umat
Chong Sheue Lih
Siti Zamratol-Mai Sarah Mukari
Department of Audiology & Speech Sciences
Faculty of Allied Health Sciences
Universiti Kebangsaan Malaysia
Jalan Raja Muda Abdul Aziz
50300 Kuala Lumpur

Corresponding author: Cila Umat
Email address: cila@medic.ukm.my
Tel: 603-26914230 / 92895005; Fax: 603-26986039

Received: April 2009
Accepted for publication: February 2010