International Journal of Audiology

Development of the Mandarin pediatric speech intelligibility (MPSI) test
Yun Zheng a; Sigfrid D. Soli b; Kai Wang a; Juan Meng a; Zhaoli Meng a; Ke Xu a; Yong Tao a
a Hearing Center, Department of Otorhinolaryngology, Head and Neck Surgery, West China Hospital, Sichuan University, China
b Department of Human Communication Sciences and Devices, House Ear Institute, Los Angeles, USA

First Published: October 2009

To cite this Article Zheng, Yun, Soli, Sigfrid D., Wang, Kai, Meng, Juan, Meng, Zhaoli, Xu, Ke and Tao, Yong (2009) 'Development of the Mandarin pediatric speech intelligibility (MPSI) test', International Journal of Audiology, 48:10, 718 — 728
To link to this Article: DOI: 10.1080/14992020902902658
URL: http://dx.doi.org/10.1080/14992020902902658

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Development of the Mandarin pediatric speech intelligibility (MPSI) test

Abstract
The objective of this research was to create a Mandarin closed-set sentence recognition test based on the English pediatric speech intelligibility (PSI) test (Jerger & Jerger, 1984) for evaluation of speech perception in children as young as three years of age. Developmentally normal children (N = 93), 3-6 years of age, were administered the Mandarin PSI (MPSI) via a computer-controlled protocol. Perfect performance was observed for all children in quiet and at +10 and +5 dB signal-to-noise ratios (SNRs). Significant age and developmental trends were seen for the more difficult SNRs, 0 dB, −5 dB, and −10 dB, with 75% of 5-6 year olds reaching the most difficult SNR. Children who reached each of the more difficult SNRs, regardless of age, exhibited the same pattern of performance on all easier conditions, indicating that the final SNR achieved, rather than percent correct scores, may be a better descriptor of performance. The MPSI comprises part of a hierarchical assessment battery for pediatric speech perception for evaluation of intervention alternatives for Mandarin-speaking children with hearing impairment.

Key Words
Pediatric speech intelligibility
Mandarin
Hierarchical speech perception test
battery
Speech perception

Abbreviations
ESP: Early speech perception test
HINT-C: Hearing in noise test for children
IT-MAIS: Infant-toddler meaningful auditory integration scale
MESP: Mandarin early speech perception test
MPSI: Mandarin pediatric speech intelligibility test
PSI: Pediatric speech intelligibility test
SNR: Signal/noise ratio

Children with hearing impairment are being identified early in life as a result of infant hearing screening programs. Consequently, there is a need for assessment tools to evaluate the success of the various intervention alternatives used to treat these very young children. These tools must capture the rapid development of speech and language that follows the onset of hearing. Eisenberg et al (2006) have advocated the use of a hierarchically structured assessment battery in evaluating early auditory awareness and subsequent development of speech perception in children from birth through adulthood. They have found that a hierarchical battery can effectively document the rapid growth in speech perception during the early stages of life in normally-hearing children, and in deaf children who receive cochlear implants (CIs) around two years of age. Their findings also suggest that with early intervention the developmental trajectory for speech perception in cochlear implant recipients may approach that of normally-hearing children of the same age. These findings define a benchmark against which the success of early intervention programs elsewhere can be compared.

In China more than 4000 children have received CIs; but only a few reports of their performance are available, in large part because of the limited availability of Mandarin language tools for assessment of speech perception in young children. Without
objective information about the performance of these children, it is difficult for clinicians and parents to identify (re)habilitation outcomes, follow up the child’s progress, and make important decisions necessary to achieve the best (re)habilitation results.

The English hierarchical speech perception test battery used by Eisenberg et al (2006) to evaluate preschool children who have received CIs includes four main tests: (1) the infant-toddler meaningful auditory integration scale (IT-MAIS) and the meaningful auditory integration scale (MAIS) (Robbins et al, 1991; Zimmerman-Phillips et al, 2000), (2) the early speech perception test (ESP) (Moog & Geers, 1990; Geers, 1994), (3) the pediatric speech intelligibility test (PSI) (Jerger & Jerger, 1984; Eisenberg & Dirks, 1995), and (4) the pediatric hearing in noise test for children (HINT-C) (Nilsson et al, 1994; Gelnert et al, 1995). A major goal of the current research is to complete the development of a set of Mandarin speech perception tests that will comprise a similar hierarchical assessment battery for use in China.

The IT-MAIS/MAIS is a structured interview with parents to document early pre-lingual auditory development including detection, recognition, and discrimination of sounds in young children during their first two years of hearing (Robbins et al, 1991; Zimmerman-Phillips et al, 2000). The ESP is a closed-set measure of simple speech perception using words for children who are unable to perform open-set measures. Normally-hearing children as young as two years can usually be tested with the ESP (Moog & Geers, 1990; Geers, 1994). It assesses the first evidence of speech perception during normal development or during development after early intervention with CIs or hearing aids. Open-set word recognition tests such as the lexical neighborhood test (LNT) and the monosyllabic LNT (MLNT) (Kirk et al, 1995; Kirk et al, 1997) may also be used; although Eisenberg et al (2006) have not included these measures in the English hierarchy. The PSI test (Jerger & Jerger, 1984; Eisenberg & Dirks, 1995) is a simple closed-set word and sentence perception test given in quiet and in the presence of competing speech for children who achieve high scores on the ESP and are able to perform closed-set sentence recognition in quiet and in noise. Normally-hearing children as young as three years can usually be tested with the PSI. The HINT-C (Nilsson et al, 1994; Gelnert et al, 1995; Vermiglio, 2008) is an adaptive open-set sentence test given in quiet or in noise. This test can usually be performed by normally-hearing children six years of age and older.

A hierarchical speech perception test battery is also critical for assessment of Mandarin-speaking preschool children. A standardized Mandarin version of IT-MAIS/MAIS has been developed (Zheng et al, 2009) by translating the English version into Mandarin. A simple closed-set Mandarin word recognition test styled after the ESP, the Mandarin early speech perception test (MESP), has also been developed and is currently used for evaluation of children with CIs and hearing aids (Zheng et al, in press). Likewise, an open-set Mandarin sentence recognition, the Mandarin HINT-C has also been developed (Wong et al, 2005). Thus, a closed-set sentence recognition test that can be administered in quiet or in noise is needed to complete a Mandarin speech perception hierarchical test battery. The current research reports the development of a closed-set sentence recognition test in Mandarin with many of the same features as PSI. Because of its similarity to the PSI, this test has been named the Mandarin pediatric speech intelligibility (MPSI) test.

The English PSI consists of a monosyllabic word test and a sentence test with ten test sentences and twenty competition sentences (Jerger et al, 1980, 1981; Jerger & Jerger, 1984). In the sentence test the child is asked to point to the corresponding picture when hearing a sentence either in quiet or in the presence of a competing sentence presented at different signal/noise ratios (SNRs). The tester ensures that the child knows all the sentences before the test starts by asking the child to ‘read’ each picture plate, i.e. to say the sentence depicted in each picture plate. The child is prompted and practiced until all picture plates can be read equally well. This ensures that the child is familiar with all of the response alternatives in the closed-set recognition task. A fixed pairing between each test sentence and competition sentence is used, and test sentences are presented in either of two formats: Format I (with carrier phrase) for younger children, and Format II (without carrier phrase) for older children. The ten test sentences are divided into two sets of five, a single set of five sentences is often repeated a second time for each test condition.

The primary purpose of the current study was to develop a closed-set Mandarin sentence recognition test, the MPSI, styled after the English PSI. The subjects in the current study were developmentally normal children who passed transient evoked otoacoustic emission (TEOAE) hearing screening in both ears. The performance of these subjects, ranging in age from 3–6 years, provides the reference necessary for comparison and interpretation of results from hearing impaired children; since the primary intended use of the MPSI is for evaluation of young hearing impaired children with CIs or hearing aids in a clinical setting. The MPSI, together with the other Mandarin speech perception assessment tools that currently exist, or are under development, will form a Mandarin hierarchical assessment battery for preschool children.

The MPSI materials are recorded in Putonghua, the standard and official Mandarin dialect in China. The majority of Chinese understand Putonghua, although young children throughout China are likely to be exposed both to Putonghua and to the Mandarin dialect of their region. The regional dialect of Mandarin in Sichuan Province, where the current research was conducted, is Sichuanhua. This dialect differs from Putonghua primarily in the makeup of its lexical tones; although there are also differences in the duration and nasalization of vowels (Norman, 1988; Ramsey, 1987; Wang et al, 2006). Zheng et al (submitted) found small but statistically reliable differences between children primarily exposed to Putonghua and children primarily exposed to Sichuanhua on the MESP subtest that assesses tone perception. Based on these results with the MESP, young developmentally normal children who are exposed primarily to Sichuanhua may perform differently on the MPSI than children who are exposed primarily to Putonghua.

The first section of this paper describes the development of the test items, test structure, and test administration procedures. Next, the MPSI test results for a group of developmentally normal children 3–6 years of age are reported. Again, these results can serve as a reference to compare the development of speech perception in hearing-impaired children with CIs and hearing aids with that of developmentally normal children. Following this, the effects on performance of early exposure to Putonghua are examined. Finally, the interpretation and clinical significance of MPSI test results are discussed.
Methods

Test structure
The approach to the development of the sentences and sentence lists for the MPSI test was similar to that used in the development of the PSI; although some differences should be noted. These differences were intended to serve as refinements using current technology, without changing the overall closed-set sentence recognition structure or the test protocol.

In the PSI test, there are two formats for the test sentences: Format I with a carrier and Format II without a carrier. In the MPSI test, only Format II is used. All children in the current study could be tested with this format, eliminating the need for the simpler Format I and reducing the time required to complete the test.

The pairing of the target and competition sentences is fixed in the PSI test, presumably because computerized test administration that would allow random pairings was not available at the time the PSI was developed. The MPSI test is computer administered, allowing randomized pairing. Such pairings are preferable because they prevent the subject from incidentally learning speech cues in the competition sentences as an aid to recognition of the pairing rather than the test sentence. This type of incidental learning would confound results, especially in test conditions with negative SNRs when the level of the competition sentence is greater than that of the test sentence.

In the PSI test, there are five test sentences per set with three different subjects. Two pairs of test sentences have the same subject while the remaining test sentence has a unique subject. A child could recognize the sentence with the unique subject by hearing only the subject, while the other sentences could only be recognized by hearing the entire sentence. The MPSI uses three pairs of sentences. Each pair has the same subject with different verbs and objects, so that none of the sentences can be recognized only by their subject.

One test sentence set is usually presented twice in each test condition when a child is tested with the PSI. When another child is tested, or the same child is retested, the other test sentence set may be used. This approach assumes the sentence sets are equivalent, although evidence of their equivalence has not been established. The MPSI uses the same two test sentence sets, presented once each, in each test condition. Thus, all test conditions for all children are based on the same sentence sets, eliminating the need to establish the equivalence of the sentence sets.

Test materials
Three criteria were used in selecting potential test sentences. First, the words and the entire sentences comprised of these words had to be within the vocabularies of a three-year old developmentally normal child. Second, the sentences had to be six to seven syllables, which is consistent with the length of the Format II test sentences in the PSI test.

A total of 50 sentences were created using the above criteria. These sentences were given in written form to three preschool teachers familiar with the language development of young children. The teachers were asked to judge whether a three-year-old child would understand each sentence and were given three choices, ‘yes,’ ‘maybe,’ and ‘no.’ Of the 50 sentences, 39 were marked ‘yes’ by all three teachers. From these sentences, 15 pairs with the same subjects and different verbs and objects were selected for further evaluation. The selected sentences were grouped in five lists of six sentences with three same-subject pairs in each list.

The sentences were illustrated by a professional artist, and the five sentence lists were administered to 15 developmentally normal children between the ages of three and four years. The child was first asked to ‘read’ each picture, i.e. to describe the picture with a sentence. Next, the sentences were presented live voice by the tester, and the child was asked to recognize the sentence by pointed to the correct illustration. The results of this testing procedure were used to select 14 sentences—two practice sentences and two sets of three same-subject pairs of test sentences—to comprise the MPSI test materials. The English and Chinese text of these sentences is given in Table 1. The pairs of sentences with the same subjects are listed sequentially for illustrative purposes.

Preparation of picture plates
Illustrations of the 14 sentences selected for the practice set and the two test sets were grouped into three picture plates—a

<table>
<thead>
<tr>
<th>Chinese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>小猫在捉老鼠</td>
<td>A cat is chasing a mouse</td>
</tr>
<tr>
<td>小猫在捉蝴蝶</td>
<td>A cat is chasing a butterfly</td>
</tr>
<tr>
<td>小猴子在爬树</td>
<td>A monkey is climbing a tree</td>
</tr>
<tr>
<td>小猴子在吃西瓜</td>
<td>A monkey is eating a watermelon</td>
</tr>
<tr>
<td>小白兔在踢球</td>
<td>A rabbit is kicking a ball</td>
</tr>
<tr>
<td>小白兔在穿鞋</td>
<td>A rabbit is putting on its shoe</td>
</tr>
<tr>
<td>小鸭在坐飞机</td>
<td>A duck is sitting in an airplane</td>
</tr>
<tr>
<td>小鸭在打电话</td>
<td>A duck in making a phone call</td>
</tr>
<tr>
<td>小老虎在洗手</td>
<td>A tiger is washing its hands</td>
</tr>
<tr>
<td>小老虎在梳头</td>
<td>A tiger is combing its hair</td>
</tr>
<tr>
<td>大象在滑滑梯</td>
<td>An elephant is sliding down a sliding board</td>
</tr>
<tr>
<td>大象在吃香蕉</td>
<td>An elephant is eating bananas</td>
</tr>
<tr>
<td>小熊猫在吃竹子</td>
<td>A panda is eating bamboo</td>
</tr>
<tr>
<td>小熊猫在洗花</td>
<td>A panda is watering flowers</td>
</tr>
</tbody>
</table>
two-picture plate for the practice sentences and two six-picture plates for the test sentences. The picture plate for the second set of test sentences is shown in Figure 1.

**Development of the competition sentences**

Three criteria were used to develop the twelve competition sentences. First, each sentence had to be longer than the longest target sentence. The longest test sentence has seven syllables, so all competition sentences have eight syllables. Second, the subjects, verbs, and objects in the competition sentences had to differ from those in the test sentences. Third, all the key words in the competition sentences had to be within the vocabulary of developmentally normal 3-4 year old children. The selected competition sentences were rated by three preschool teachers using the same procedure as for the test sentences. All teachers responded ‘yes’ or ‘maybe’ to the question of whether the competition sentences could be understood by children in this age range. (One teacher rated two of the sentences ‘maybe.’) All other sentences were rated ‘yes’ by all teachers.) The Chinese and English text of these sentences is given in Table 2.

**Recording the MPSI sentences**

The MPSI practice, test, and competition sentences were recorded at the House Ear Institute using the procedures described by Nilsson et al (1994). The talker for the recordings is a native of Beijing with a professionally trained voice who speaks Putonghua as a news reader for a local Mandarin language radio station. He is the same talker used for the Mandarin HINT and the MESP recordings (Wong, et al 2005; Zheng et al, in press). The recordings were edited into individual sentence files, the A-weighted root-mean-square (RMS) amplitude of each sentence was calculated, and the sentences were scaled to the same RMS level.

**Test administration and scoring**

A software program was written for administration and automated scoring of the MPSI test. The sentence set and the sentences within the set are selected randomly by the program and presented via loudspeaker. Each target sentence is randomly paired with one of the competition sentences at the specified SNR for the condition. A picture-pointing task with printed picture plates is used to obtain the child’s responses. Responses to the MPSI test are entered into the software after each trial and are automatically scored at the end of the sentence set as the percent of sentences in the set that were correctly recognized. Both sentence sets are administered and scored in this manner for each test condition. The final score for the condition is the percent correct sentence recognition for all 12 sentences.

Testing started with the easiest condition and progresses through increasingly difficult conditions from +10 dB SNR to −10 dB SNR in 5-dB steps until the subject failed to obtain a score significantly above chance performance at the p = 0.05 level. Chance performance is 2 of 12 correct, or 16.7%, with a score of five correct (41.7%) or more significantly above chance. If the score for a condition is above chance for that condition,

---

**Table 2. The MPSI competition sentences.**

<table>
<thead>
<tr>
<th>Chinese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>小老鼠在玩搭积木</td>
<td>A mouse is piling up its building blocks</td>
</tr>
<tr>
<td>小绵羊在擦玻璃窗</td>
<td>A sheep is cleaning the glass windows</td>
</tr>
<tr>
<td>大公鸡在喝果汁水</td>
<td>A cock is drinking fruit juice</td>
</tr>
<tr>
<td>小鸟在教大家唱歌</td>
<td>A bird is teaching the others singing</td>
</tr>
<tr>
<td>小绵羊在开电冰箱</td>
<td>A sheep is opening the refrigerator</td>
</tr>
<tr>
<td>小鸟在喂孩子虫子</td>
<td>A bird is feeding worms to the baby bird</td>
</tr>
<tr>
<td>小牛在买夹心饼干</td>
<td>A cow is buying sandwich biscuits</td>
</tr>
<tr>
<td>小熊在做切片面包</td>
<td>A bear is making sliced bread</td>
</tr>
<tr>
<td>小乌龟在上幼儿园</td>
<td>A tortoise is going to the kindergarten</td>
</tr>
<tr>
<td>小马在看小羊跳舞</td>
<td>A horse is watching a sheep dancing</td>
</tr>
<tr>
<td>小青蛙在看电影</td>
<td>A frog is watching a cartoon movie</td>
</tr>
<tr>
<td>小牛在买夏凉包</td>
<td>A cow is buying red bean buns</td>
</tr>
</tbody>
</table>
the child progresses to the next condition. When a condition is reached where the child’s score does not significantly exceed chance performance, the immediately preceding condition is defined as the final SNR reached by the child and the score for that condition is recorded.

Each child was tested using the software program described above. The computer screen and keyboard were oriented away from the child’s view so that only the appropriate picture plate was visible to the child. A pair of small powered loudspeakers was connected to the notebook computer’s audio output with one in front of the child and one behind the child at equal distances from the center of the child’s head. The output levels of the loudspeakers were matched and calibrated at the position of the center of the head. The output level of the front loudspeaker used for presentation of the test sentences was adjusted to a comfortable listening level by the tester before the child was seated. The various SNRs that define the test conditions were achieved under software control by adjustment of the competition level from the rear loudspeaker. All testing was done in a quiet room at the preschool where the children were students.

The child was seated so the center of their head was midway between the two loudspeakers during test administration. The child was asked to read the pictures on the picture plates first, and then to point to the corresponding picture when hearing the sentences with live voice. If the child could not read a picture correctly, the tester would practice the child on the test sentence until the child had learned all the key words of the sentence. This step was included to ensure that all the closed-set response alternatives were equally familiar to the child. The practice test was administered first in this manner to familiarize the child with the task, followed by the two sets of test sentences. The order of the two test sets was randomized by the software, as was the order of the sentences within each set and the pairing of the test and competition sentences. After the practice test, the quiet condition was administered first, followed by the tests with competing speech at different SNRs that progressed from easiest (+10 dB SNR) to most difficult (−10 dB SNR). All responses were recorded automatically by the tester in the software.

When the MPSI is used to evaluate hearing-impaired children the order of test conditions from easiest to most difficult will be used to find the most difficult condition at which the child can perform above chance. The same order of test conditions was used with the developmentally normal children in the current study to provide a reference for comparison with the results from hearing-impaired children. Use of a randomized or counterbalanced order of test conditions in the current study would confound such comparisons. Effects of test order on performance can occur because of the large differences in difficulty between the MPSI test conditions, with SNRs ranging over 20 dB. Thus, the strategy was to control these effects by fixing the test order for all subjects to be the same as for the hearing-impaired child.

The child was allowed three opportunities to recognize the first sentence in each new condition. The subsequent sentences in each condition were presented once.

**Dialect exposure checklist**

Previous research on the development of the MESP (Zheng et al., 2009) revealed that the amount of early exposure to Putonghua can affect the speech perception of developmentally normal children, especially in tasks that rely entirely on tone perception. Although the sentences in the MPSI are composed of differing segmental and tonal cues, it is still possible that the amount of Putonghua exposure may affect performance. The same simple dialect exposure checklist as that used by Zheng et al (in press) was used in the current research to assess each child’s exposure to Mandarin. This checklist was given to the caregivers and teachers of the children. It consists of questions about the child’s dialect exposure at the preschool and at home, as observed by caregivers and teachers. The checklist is described in the Appendix. Answers to each question were given in terms of percent Putonghua exposure and usage. Item responses were based entirely on the judgments of the caregivers and teachers, none of whom had been instructed or practiced in the evaluation of dialect exposure. Thus, the primary value of the data from the checklist is to determine only the approximate amount of Putonghua exposure for each child.

**Subjects**

Subjects ranging in age from 3–6 years were recruited from two preschools in urban Chengdu. Letters were sent to the parents of all children in this age range from these preschools inviting them to participate. The developmental status of each respondent who accepted the invitation was determined from their teacher’s responses to a checklist for language and cognitive development. The teacher most familiar with each child was asked to rate the child’s language development and receptive language as ‘delayed,’ ‘normal,’ or ‘advanced.’ All of the potential subjects from both preschools received ‘normal’ ratings. These children received a TEOAE hearing screening at the preschool (Madsen Accuscreen). Children who passed the hearing screening for both ears were selected to participate in the study. A total of 93 children were chosen in this manner. The age and gender distribution of the sample is summarized in Table 3.

**Results**

All children could be tested with the computer-administered protocol for the MPSI test. The time required for testing was approximately 20 minutes per child. Since the children were divided into three age groups, [3–4), [4–5), and [5–6) (‘[’ is defined as ‘greater than or equal to,’ and ‘)’ is defined as ‘less than’), and since the test conditions included quiet and five competition conditions with different SNRs from +10 to −10 dB SNR in 5-dB steps, MPSI test results are reported as a function of age group and test condition. Analyses of the dialect exposure checklist results are reported separately.

**Table 3. Age and gender distribution for the sample.**

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>[3–4)</th>
<th>[4–5)</th>
<th>[5–6)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>17</td>
<td>14</td>
<td>16</td>
<td>47</td>
</tr>
<tr>
<td>Female</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>46</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>29</td>
<td>32</td>
<td>93</td>
</tr>
</tbody>
</table>

\[\text{['] = greater than or equal to.}\]

\[\text{[') = less than.}\]
Age and SNR effects

Only those scores significantly above chance were included in the data analyses; i.e., if a child scored significantly above chance in the first three SNR conditions and not in the fourth condition, only the first three scores were included. The score for each condition was based on the responses to all six sentences in the list, including the first sentence which may have been repeated. The occurrence of repetitions was not recorded by the software; although the tester observed that most of the repetitions occurred with the youngest children in the most difficult SNR conditions. Thus, the effect of repetition on scores is minimal, because repetitions occurred primarily in the most difficult conditions where the child was unable to score significantly above chance. Scores from these conditions are not used.

Table 4 displays the MPSI results in two ways: the percent of children performing above chance, and the average percent correct score at each SNR in each age group for those children scoring significantly above chance. The top portion of Table 4 shows that all children performed significantly above chance in the first three conditions: quiet, +10 dB SNR, and +5 dB SNR. However, the percent of children who performed above chance dropped consistently and at different rates for each age group as the test conditions became more difficult. For the [3–4] year olds, only 3.1% of the 32 children in this age group reached the −10 dB SNR condition, and 34.4% of the 32 children reached −5 dB SNR. The two older groups showed similar declining proportions, with many more of the [5–6] year old children reaching the most difficult conditions. A chi-squared analysis revealed that the distribution of children reaching each of the three most difficult conditions differed significantly across age groups, $\chi^2(4)=45.83, p<0.0001$.

The second part of Table 4 gives the average percent correct scores in each condition for the children in each age group who reached that condition. Within each age group, the percent correct scores decrease as the test condition gets more difficult, as expected. Again, the percent correct performance of the younger children decreased more rapidly than that of the older children.

Although it is possible to perform statistical tests that compare performance of children in each age group who reach each test condition, it is more informative to separate the sample into those subjects who reached the −10 dB SNR condition regardless of age, those subjects who reached the −5 dB SNR condition regardless of age, and those subjects who reached the 0 dB SNR condition regardless of age. This approach uses performance on the MPSI rather than age, at best an indirect measure of developmental status, as a grouping variable. Within the three groups of subjects defined in this manner, the effects of age and SNR on MPSI performance may be compared. Since only one child in the [3–4] year group reached the −10 dB SNR condition, comparisons for this condition are made for only the oldest two groups. Likewise, since none of the children in the [5–6] year old group reached 0 dB as their final condition, the age comparisons within the group reaching 0 dB SNR are made only for the two youngest groups. Children of all ages reached the −5 dB SNR condition in sufficient numbers, so comparisons between all three age groups can be made.

Figure 2 shows the average performance for each age as a function of SNR for the groups of subjects reaching each of the three most difficult SNR conditions. Panel (a) displays the means for children reaching the −10 dB SNR condition. Panel (b) displays the means for children reaching the −5 dB SNR condition; and Panel (c) displays the means for children reaching the 0 dB SNR condition. The pattern of mean performance in each panel shows several common characteristics. Performance among subjects who reach the same final condition is generally quite similar regardless of age. In addition, in Panels (b) and (c) performance remains high (averaging above 90%) until the final condition is reached, and then drops off precipitously.

Separate statistical analyses of the data for each group revealed no significant age effects for any of the groups, regardless of the final SNR reached (for −10 dB SNR: $F(1, 30)=0.58, p=0.811$; for −5 dB SNR: $F(1, 22)=0.42, p=0.522$; and for 0 dB SNR: $F(1, 22)=0.42, p=0.522$). All of the SNR effects, however, were statistical significance (for −10 dB SNR: $F(4, 27)=20.27, p<0.0001$; for −5 dB SNR: $F(3, 28)=38.58, p<0.0001$; and for 0 dB SNR: $F(2, 21)=14.24, p<0.0001$).

### Table 4. Summary of the MPSI test results.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Quiet</th>
<th>+10 dB</th>
<th>+5 dB</th>
<th>0 dB</th>
<th>−5 dB</th>
<th>−10 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>[3–4]</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>90.6</td>
<td>34.4</td>
<td>3.1</td>
</tr>
<tr>
<td>[4–5]</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>79.3</td>
<td>27.6</td>
</tr>
<tr>
<td>[5–6]</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>96.7</td>
<td>75.0</td>
</tr>
</tbody>
</table>

**Average percent correct score at each SNR**

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>3–4</th>
<th>+10 dB</th>
<th>+5 dB</th>
<th>0 dB</th>
<th>−5 dB</th>
<th>−10 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>[3–4]</td>
<td>100</td>
<td>96.6</td>
<td>94.1</td>
<td>79.9</td>
<td>63.6</td>
<td>42.0</td>
</tr>
<tr>
<td>[4–5]</td>
<td>100</td>
<td>96.7</td>
<td>95.4</td>
<td>88.9</td>
<td>69.2</td>
<td>58.4</td>
</tr>
<tr>
<td>[5–6]</td>
<td>100</td>
<td>99.8</td>
<td>97.9</td>
<td>90.2</td>
<td>80.7</td>
<td>66.7</td>
</tr>
</tbody>
</table>

**Percent Mandarin exposure for subjects performing above chance at each S/N**

<table>
<thead>
<tr>
<th>All ages</th>
<th>□</th>
<th>□</th>
<th>□</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

\[\text{\textsuperscript{\dag}}\text{=greater than or equal to.}\]

\[\text{\textsuperscript{\dag\dag}}\text{=less than.}\]
To summarize, the preceding analyses indicate that children who reach the same final SNR condition on the MPSI test have almost identical performance on each of the preceding easier SNR test condition, regardless of age. However, younger children are not as likely to reach the more difficult conditions on the MPSI test, as shown from the chi-squared analysis. This pattern of results suggests that the final SNR condition a developmentally normal child reaches may be the more informative than the performance achieved on the preceding SNR conditions. Once a child reaches a particular SNR condition, their performance accuracy on the preceding SNR conditions is about the same, regardless of age. This pattern of results and related observations may, of course, differ for hearing-impaired or developmentally delayed children.

The pattern of results shown in the three panels of Figure 2 is summarized in Figure 3. This figure displays MPSI test performance at each SNR condition averaged across all ages for the three groups of subjects reaching the most difficult SNR conditions. Note in Figure 3 that all three groups exhibit similar patterns of performance in the easiest +10 and +5 dB SNR conditions. The group that reached 0 dB SNR shows a large drop in performance when this condition is reached; while the remaining two groups continue their pattern of gradually decreasing performance until the group that reached −5 dB SNR exhibits a large drop when this condition is reached. The group reaching the −10 dB SNR condition continues their gradual decrease, and then drops by about 20% when the final condition is reached.

These results show consistent age differences in the proportion of children reaching each of the most difficult SNR conditions. However, when the children who reach each final SNR condition are compared across age, there are no significant age differences in percent correct performance on the MPSI test. A consistent pattern of performance at all SNR conditions is also seen when subjects are grouped according to the final SNR condition they reach, regardless of their age. Thus, the final SNR condition reached by the subject may be a better indication of their speech perception abilities than age, or the percent correct scores they receive in the easier SNR test conditions.

Dialect exposure effects

Dialect exposure checklists were given to each child’s teacher and caregiver. A total of 94% of the checklists were returned. Responses to the checklist items (see Appendix) were given as the percentage of time Putonghua was used in spoken communications to the child by the teacher and caregiver, as well as by the child to the teacher, caregiver, and peers. Results for the checklist are shown in Table 5.

The results show that overall percent exposure to Putonghua, defined as the unweighted mean of the percent exposures the child encountered in the preschool and home environments, averaged from 48.1% to 63.7% across age groups. The standard deviations associated with these means were large, indicating a wide range of Putonghua exposure within each age group. Putonghua exposure in the preschool environment was substantially higher than in the home environment. Most teacher-child communications used Putonghua. Fewer than half of caregivers reported Putonghua as their primary dialect, and most caregiver-child communications did not use Putonghua. The average age at first Putonghua exposure ranged from approximately 0.8–1.6 years, and the first use of Putonghua occurred less than a year afterwards on average.
Table 5. Summary of the language checklist results. Exposure means and standard deviations express the percentage of time teachers and caregivers used or observed the use of Putonghua in communications with the child. The caregiver dialect expresses the percent of caregivers whose primary dialect was Putonghua.

<table>
<thead>
<tr>
<th>Age</th>
<th>N</th>
<th>Summary statistics</th>
<th>Overall exposure</th>
<th>Preschool environment</th>
<th>Home environment</th>
<th>Age of first exposure and first use (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>With teacher</td>
<td>With caregiver</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>To child</td>
<td>By child</td>
<td>To child</td>
</tr>
<tr>
<td>[3-4)</td>
<td>29</td>
<td>Mean</td>
<td>56.4%</td>
<td>99.7%</td>
<td>84.0%</td>
<td>49.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std. Dev.</td>
<td>29.1%</td>
<td>1.9%</td>
<td>29.5%</td>
<td>48.7%</td>
</tr>
<tr>
<td>[4-5)</td>
<td>30</td>
<td>Mean</td>
<td>63.7%</td>
<td>93.7%</td>
<td>92.5%</td>
<td>83.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std. Dev.</td>
<td>22.2%</td>
<td>21.4%</td>
<td>21.8%</td>
<td>28.0%</td>
</tr>
<tr>
<td>[5-6)</td>
<td>29</td>
<td>Mean</td>
<td>48.1%</td>
<td>97.2%</td>
<td>43.4%</td>
<td>40.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std. Dev.</td>
<td>27.8%</td>
<td>11.3%</td>
<td>45.9%</td>
<td>43.3%</td>
</tr>
</tbody>
</table>

'\geq' = greater than or equal to.
'\lt' = less than.
Although the dialect exposure checklist provides inexact information from untrained informants, it nonetheless reveals a large range of experience with Putonghua among the subjects in the study. A similar range of experience with Putonghua is to be expected when the MPSI is used to evaluate hearing-impaired children. Thus, it is important to determine whether dialect exposure is likely to affect MPSI results. To address this question the overall percentage of Putonghua exposure, as defined above, was tabulated for all subjects, regardless of age, who reached each final test condition.

Statistical tests comparing Putonghua exposure for children reaching the $-10$ dB SNR test condition with that for children reaching the $0$ dB SNR test condition and the $-5$ dB SNR test condition revealed nonsignificant differences in exposure between these groups, $t(60) = 0.57, p = 0.57$, and $t(46) = 0.68, p = 0.50$, respectively.

Percent Putonghua exposure was also correlated with MPSI scores within each of the three groups defined by the final SNR test condition reached. For Putonghua exposure to have an effect on performance for the MPSI test, correlations between the percent of exposure should achieve statistical significance, and a meaningful amount of variance or variability in MPSI test scores should be predictable from the amount of exposure for each child. None of these correlations achieved statistical significance, and none accounted for more than 5% of the variance in MPSI scores ($r = 0.21, 0.03$, and $0.12$ for $-10$ dB, $-5$ dB, and $0$ dB SNR, respectively). In short, dialect exposure does not appear to affect MPSI test results in developmentally normal children significantly.

**Discussion**

The primary objective of this research was to develop a closed-set Mandarin sentence recognition test suitable for use with young preschool children who can be to be evaluated with such materials in quiet and in noise. The need exists for such a test within the Mandarin hierarchical assessment battery to evaluate the development of speech perception in children with hearing impairment who are treated with CIs or hearing aids. The Mandarin battery is meant to parallel the English battery proposed by Eisenberg et al. (2006) in which the PSI serves as the closed-set sentence recognition test. The MPSI has the same test structure and protocol as the PSI; although refinements have been made using current technology. The current research establishes the applicability of the MPSI within the Mandarin hierarchy for assessment of closed-set sentence perception skills in children as young as three years. This research also provides reference data from developmentally normal children for other purposes it may possible to establish a basal SNR for different age groups as a more efficient means of testing.

The results also showed that there were no significant Mandarin dialect (Putonghua) exposure effects on MPSI test results. This finding differs from the results of Zheng et al. (in press) showing small but significant dialect exposure effects on performance in the tone perception subtest of the MESP. This subtest is comprised of minimally contrastive monosyllabic words that differ only in tone, while the sentences in the MPSI test differ in their segmental and tonal makeup. The absence of dialect effects suggests that the segmental and lexical tone differences between Putonghua and Sichuanhua are not substantial enough in the MPSI materials to affect the results because of the redundancy of information in the sentences. Thus, the MESP tone subtest may be more sensitive to dialect exposure differences than the MPSI sentence test.

The intended use of the MPSI test is to evaluate children with hearing aids, cochlear implants, and other prosthetic devices to determine the benefits of these interventions. The current results for developmentally normal children do not guarantee that all features of the MPSI test will be equally usable with children who are hearing-impaired and/or developmentally delayed. For example, additional training and familiarization with the sentences and picture plates may be needed to ensure that the

---

726
test items are within the child’s vocabulary. It is also important to note other potential limitations of the current research. The sample in the current study was comprised of a limited number of children from two urban preschools in one city. Moreover, teacher screening was used as the sole criterion for normal language development. Likewise, the dialect exposure characteristics of the sample were determined by ratings from teachers and caregivers who were familiar with the child, but inexperienced as raters. These considerations may limit the ability to generalize the current results. Additional research with larger and better defined groups of developmentally normal children is required to reveal the extent of such limitations.

The MPSI test is currently being used to evaluate children with hearing impairment in China. Results from this clinical research will be reported in future publications. The results from this study with developmentally normal children can serve as a reference for comparisons to the performance of children with hearing impairment who have received various early interventions. The similar structures and protocols of the PSI and the MPSI tests may make it possible to compare MPSI test results for Mandarin-speaking children to results obtained with the PSI test for English-speaking children. Such comparisons can serve to advance the audiological underpinnings of early intervention and to provide international documentation of their successes.

Acknowledgements

The authors gratefully acknowledge Daniel Freed and Andrew Vermiglio at the House Ear Institute for their assistance in recording the MPSI test materials and for development and testing of the software used to administer and automatically score the test. The Hearing Center, Department of Otorhinolaryngology, Head and Neck Surgery, at West China Hospital of Sichuan University is also gratefully acknowledged for their support of the international collaboration necessary to complete this research. Special gratitude is due to the preschool teachers who volunteered their time and opened their schools to allow the MPSI testing to take place. Finally, we thank the parents of our subjects for allowing their children to participate in this research.

This study was conducted according to the ethics procedures established by West China Hospital for testing of human subjects. Written informed parental consent was obtained for participants in the study.

Declaration of interest: This research was funded by the participant organizations, House Ear Institute and West China Hospital. The MPSI test and related computer-administered test protocols have been licensed by the participant organizations to the Asia Pacific Division of Cochlear, Ltd., Sydney, Australia, for free distribution within Asia. License fees received by the participant organizations have been used to defray the cost of the research. The authors received no payments from these fees, nor do they have any financial interests in the licensee.

References


Appendix

Dialect exposure checklist

The same dialect exposure checklist as that used by Zheng et al (in press) was given to the caregivers and teachers of the children in the current study. To summarize the description by Zheng et al, the two-part checklist consisted of questions about the child’s dialect exposure at preschool and at home. Questions were asked of the child’s teacher, and of the child’s primary caregiver, who could be a parent, grandparent, or helper.

Three questions were asked of the teacher:
1. What percent of your communication with the child uses Putonghua?
2. What percent of the child’s communication to you uses Putonghua?
3. What percent of the child’s communication with his or her peers uses Putonghua?

Five questions were asked of the caregiver:
1. What percent of your communication with the child uses Putonghua?
2. What percent of the child’s communication to you uses Putonghua?
3. What is your primary Mandarin dialect? (Putonghua, Sichuanhua, other)
4. At what age was the child first exposed to Putonghua?
5. At what age did the child start to speak Putonghua?

Answers to each question, except those about age, were given in terms of percent Putonghua and percent Sichuanhua. Responses with mention of other dialects, which occurred rarely, were excluded from consideration.