Chinese Word Identification and Sentence Intelligibility in Primary School Classrooms

Jianxin PENG, Peng JIANG

School of Physics and Optoelectronics, South China University of Technology
Guangzhou, 510640, China; e-mail: phjxpeng@163.com

The Chinese word identification and sentence intelligibility are evaluated by grades 3 and 5 students in the classrooms with different reverberation times (RTs) from three primary school under different signal-to-noise ratios (SNRs). The relationships between subjective word identification and sentence intelligibility scores and speech transmission index (STI) are analyzed. The results show that both Chinese word identification and sentence intelligibility scores for grades 3 and 5 students in the classroom increased with the increase of SNR (and STI), increased with the increase of the age of students, and decreased with the increase of RT. To achieve a 99% sentence intelligibility score, the STIs required for grades 3, grade 5 students, and adults are 0.71, 0.61, and 0.51, respectively. The required objective acoustical index determined by a certain threshold of the word identification test might be underestimated for younger children (grade 3 students) in classroom but overestimated for adults. A method based on the sentence test is more useful for speech intelligibility evaluation in classrooms than that based on the word test for different age groups. Younger children need more favorable classroom acoustical environment with a higher STI than older children and adults to achieve the optimum speech communication in the classroom.

Keywords: speech identification; reverberation time; signal-to-noise ratio; classroom; speech transmission index; children.

1. Introduction

The acoustical environments in elementary school classrooms have an important role in developing the ability of learning and cognition of children. Evidence shows that poor room acoustical environment, such as excessive noise and reverberation, reduces the speech intelligibility in a classroom, interrupts the verbal communication between teachers and children, and has a negative influence on the memory of children (HYGGE, 2003). Thus, high speech intelligibility is essential in classrooms. Background noise, reverberation time and SNR all have a critical impact on speech communication in the classroom. Many studies had investigated speech intelligibility in classrooms (BRADLEY, 1986; HOUTGAST, 1981; BRADLEY, SATO, 2008; TANG, YEUNG, 2004; BRACHMANSKI, 2008; 2012; ASTOLFI *et al.*, 2012, PENG *et al.*, 2015) and the relationship between speech intelligibility scores and objective acoustical indices, such as speech transmission index (STI), A-weighted signal-to-noise ratio (SNR(A)) and useful/detrimental sound ratio for different language (HOUTGAST, 1981; BRADLEY, 1986;

ANDERSON, KALB, 1987; BRACHMAŃSKI, 2004; PENG *et al.*, 2015).

HOUTGAST (1981) conducted a series of speech intelligibility tests in classrooms for teachers and children aged 8 to 15 years under a variety of road traffic noise conditions with reverberation times (RTs) from 0.7 s to 1.5 s. The relationship between SNR(A) and speech intelligibility scores for Dutch was reported. BRADLEY (1986) investigated speech intelligibility using the English Fairbank rhyme test in realistic elementary classrooms with RTs from 0.39 s to 1.20 s for children aged 12 to 13 years through a small loudspeaker with its directivity similar to a human's mouth. The relationship between English speech intelligibility scores and SNR(A), useful/detrimental sound ratios were founded, respectively. BRADLEY and SATO (2008) likewise performed a series of speech intelligibility tests using the Word Identification by Picture Identification test for children in grades 1, 3, and 6 in realistic elementary school classrooms. The relationship between SNR(A) and English speech intelligibility scores was also revealed. ASTOLFI *et al.* (2012) explored Italian speech intelligibility using a diagnostic

rhyme test in four elementary classrooms with different RTs for children aged 7 to 10 years old. The relationship between STI and Italian speech intelligibility scores was also built. The results showed that children in grade 2 understood fewer words than children in the higher grades when the STI was lower. PENG et al. (2015) investigated the Chinese speech intelligibility of elementary school children from grades 2, 4 and 6 with noise and reverberation in 28 active elementary classrooms in 9 different elementary schools. The relationship between Chinese speech intelligibility scores and STI was obtained.

However, a limitation of word lists for evaluating the effects of noise and reverberation on speech identification is that the test materials are not presented in a manner representative of a children's everyday speech communication in the classroom (YACULLO, HAWKINS, 1987). KITAPCI and GALBRUN (2014) revealed that there is a significant difference among the subjective word identification scores of English, Polish, Arabic and Mandarin, but no statistically significant difference among the subjective sentence intelligibility scores of these languages. There are two basic operations involved in the identification of words in sentences (STUART, 2008). One is the initial processing of acoustic-phonetic information, the other is the utilization of linguistic-situational information of speech (STUART, 2008; KALIKOW et al., 1977). If the purpose of speech intelligibility test is to assess everyday vocal communication ability in realistic classroom acoustical environments, the information obtained from isolated words will be limited. It is more important to identify and understand the meaning of a sentence than to recognize the single word. YACULLO and HAWKINS (1987) used sentence materials to examine the effects of noise (SNR = +2 and +6 dB) and RTs (0.0 and 0.8 s) on monaural speech identification in normal-hearing school-aged children (aged 8–10 years). Results revealed that the typical acoustic conditions existing in classrooms do not permit adequate speech identification. The introduction of a typical level of classroom reverberation significantly reduced speech identification in both levels of noise. OZIMEK et al. (2013) investigated the speech intelligibility for different configurations of a target signal (speech) and masker (babble noise) using a Polish sentence test. The best performance in speech intelligibility was found for the binaural mode. STUART (2008) investigated sentence intelligibility in noise by school-age children. Reception thresholds for sentences were determined in quiet and in backgrounds of competing continuous and interrupted noise. Results show that children are more adversely affected by noise and needed greater SNR in order to perform as well as adults.

Sentence materials provide more realistic listening conditions for everyday communication in classrooms. KOCIŃSKI and OZIMEK (2015) measured sentence and logatome speech intelligibility in rooms with induction loop for hearing aid users. Their results show that a sentence test is a more useful tool for speech intelligibility measurements in a room than logatome test. The studies mentioned above are based on Western languages, the results may not be suitable for children in China because Chinese is a tonal language that is different from Western languages. In this study, the word identification and sentence intelligibility were evaluated by children in three realistic primary school classrooms with different RTs under different SNRs conditions. The word identification and sentence intelligibility scores were compared and analyzed.

2. Experimental methods

2.1. Classrooms

Three rectangular classrooms were selected. Their dimensions were $9.35 \times 7.90 \times 3.70$ m^3 (Classroom A), $8.60 \times 6.15 \times 3.25$ m^3 (Classroom B), $9.35 \times 7.90 \times 3.15$ m^3 (Classroom C). There was no sound absorption treatment for classrooms A and B. Both sides of the walls besides the windows were plastered in classrooms A and B. There were two blackboards in the front of the classroom and in the back of the classroom, respectively. The ceiling was also plastered and the floor was covered with ceramic tile. There were about 45 desks in each classroom. For classroom C, mineral-fiber acoustic ceiling tiles with 1.5 cm thickness were installed on the ceiling. Except for all window curtains closed in classroom B, the other interior surfaces of classrooms A and B were the same. For each classroom, a sound source (JBL-6325P loudspeaker) was set at 1.5 m above the floor in front of the room where the teacher would normally stand. The listening position was located at the back of the classroom (see Fig. 1). The average RT, EDT and D$_{50}$
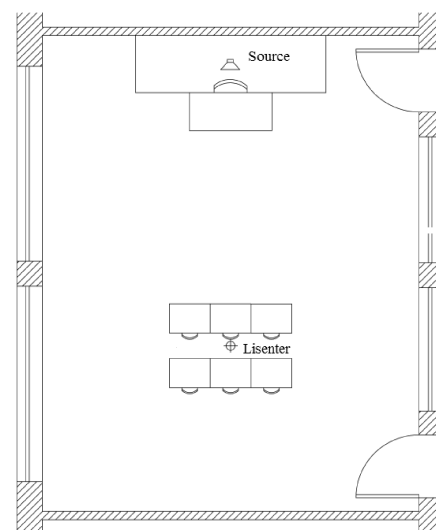

Fig. 1. Sound source and listening position
in classrooms.

in 500∼2000 Hz octave band at the listening positions in classrooms A, B and C were measured, and are shown in Table 1.

Table 1. The objective acoustical indices in three classrooms.

Classroom	EDT/s	T_{30}/s	D_{50}	STI
A	1.17	1.05	0.49	0.55
B	0.89	0.77	0.61	0.62
C	0.39	0.47	0.86	0.76

2.2. Word identification and sentence intelligibility test

The word identification test used Mandarin Chinese test word lists as specified by GB 4959 (1985). The test used 25 five-word rows of similar-sounding Chinese monosyllabic words and was similar to modified rhyme test of English. The five words in each row were randomly arranged and differed only in the initial consonant sound (for example, hao, sao, gao, zao, kao). Thirteen word lists were derived from five basic word lists. One list was used as a practice list for students, the remaining twelve word lists were used for the test. The test word in carrier sentence was "The × row reads −". The "×" stands for row number and "−" stands for a test word. Chinese Mandarin Hearing in Noise Test (MHINT) (Wong *et al.*, 2007) was used for the sentence intelligibility test. Every list contained 10 test sentences and each sentence was composed of 10 target words, for example, "这个球队终于打入决赛 (This team enters the finals at last)". All word and sentence test lists were recorded at the rate of 4.0 words per second in an anechoic chamber by using an omni-directional precision microphone at a distance of 0.5 m from the male speaker. All recordings were edited by CoolEdit Pro. For the word test signal, a ten-second interval of silence between two adjacent carrier sentences was added to allow subjects to mark the test word that they had heard on the test paper. For the sentence test, a thirty-second interval of silence between two adjacent test sentences was added to allow subjects to write the test sentence that they had heard.

Based on the average speech spectrum of test word signals and test sentence signals, the corresponding speech-shaped noises were selected for the tests. The SNR was made equal for all selected frequency bands using speech-shaped noise with a frequency spectrum equivalent to the long-term speech spectrum. The test signals recorded in an anechoic chamber and speech-shaped noise were mixed with six different SNRs (i.e. 0 dBA, 5 dBA, 10 dBA, 15 dBA, 20 dBA and ∞), then were reproduced by a JBL-LSR6325P loudspeaker with its directivity similar to the human mouth. The SNR ∞ means ambient noise only (ANO) for the test.

At this case, no speech-shaped noise was mixed with the test signals and the ambient noise was the only noise source for the test in the classroom. The speech level was set to 70 dBA in 0 dBA, 5 dBA, 10 dBA, 15 dBA SNRs and ANO case, and 75 dBA in 20 dBA SNR case. For each case, two test lists were used. During the test, speech and noise signals were recorded at each listening position in every classroom. These recordings were used to determine speech and noise levels for calculating the STI at each listening position in every test condition.

2.3. Subjects

All children were randomly selected from grades 3 (8 to 9 years old) and 5 (10 to 11 years old) students in those primary schools participated in the test, respectively. They were representative of general listening audiences. No hearing tests for children were performed before the speech identification test. All children reported that they have no known hearing problems and can speak and understand standard Mandarin Chinese. They received a few minutes of instruction prior to the tests. For each testing condition, the subjects evenly sat around listening positions in each classroom. An average of the correct word identification and sentence intelligibility scores across all subjects and two lists were obtained for each testing condition.

3. Results

The subjective Chinese word identification and sentence intelligibility at the listening position for grades 3 and 5 students was conducted in three classrooms with the different SNRs. Figure 2 shows the Chinese word identification and sentence intelligibility scores and their standard deviations (SDs) in different classrooms and different SNRs, respectively.

It can be seen from Fig. 2 that, both Chinese word identification and sentence intelligibility scores for grades 3 and 5 students in the classroom increased with the increase of SNRs and decreased with the increase of RTs. The result was in good agreement with that of existing studies for Western languages (Bradley, 1986; Yacullo, Hawkins, 1987; Bradley, Sato, 2008; Stuart, 2008; Astolfi *et al.*, 2012; Tillery *et al.*, 2012). The Chinese word identification and sentence intelligibility scores also increased with the increase of the age of students.

Figure 3 shows that the difference between word identification scores and sentence intelligibility scores under the same test condition. Under five of the 18 test conditions, the sentence intelligibility scores were lower than the word identification scores for grade 3 students. For grade 5 students, the sentence intelligibility scores were lower than the word identification scores under one test conditions only.
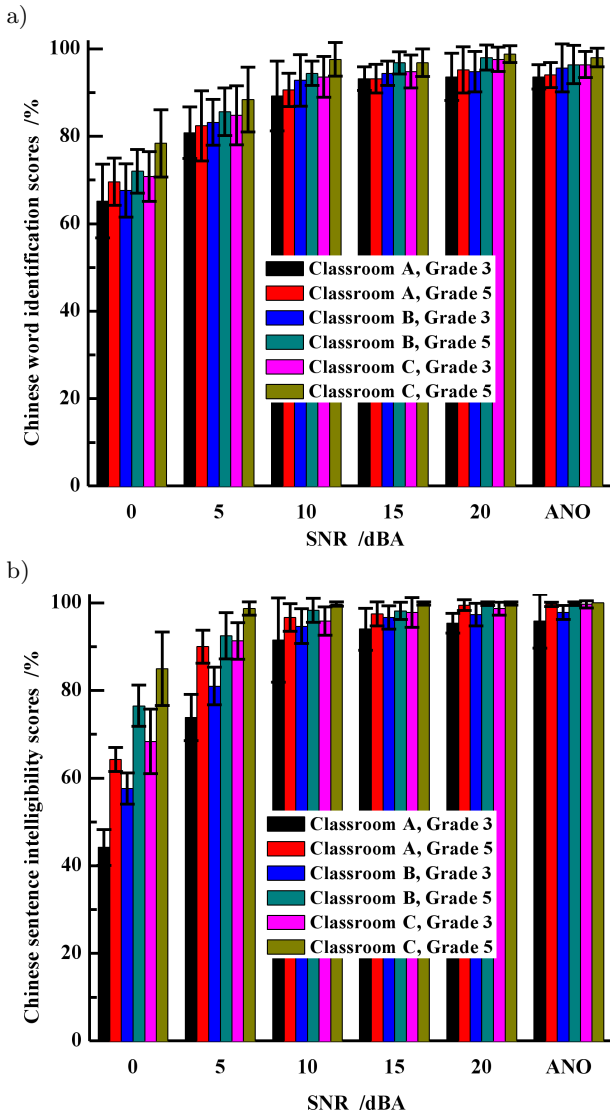
a)

b)

Fig. 2. Chinese word identification and sentence intelligibility scores and their standard deviations: a) word identification, and b) sentence intelligibility.
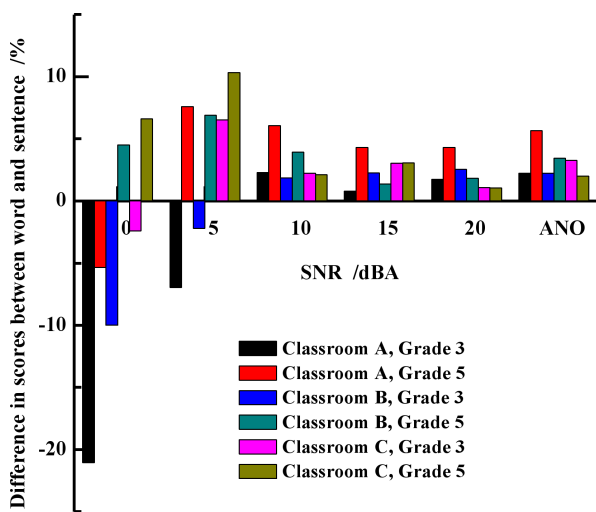

Fig. 3. The difference in scores between Chinese word identification and sentence intelligibility.

The Chinese word identification and sentence intelligibility scores for grades 3 and 5 students were plotted against the corresponding STI from different test conditions in Figs. 4 and 5. The correlation coefficients (R) and SDs of the best-fit curve to the data of Figs. 4 and 5 are shown in Table 2. To compare the difference in scores from different age groups, the re lationship between STI and Chinese word identification and sentence intelligibility scores of adults was
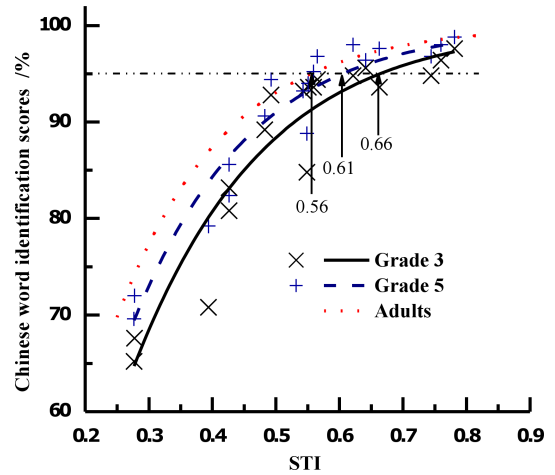

Fig. 4. The relationship between Chinese word identification scores and STI and best-fit curves, the data of adults came from the study by Peng (2008).
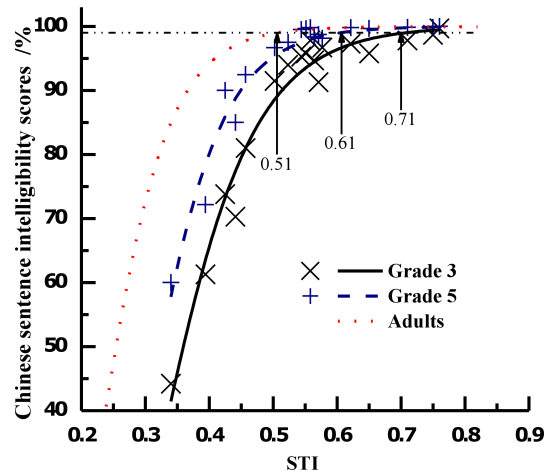

Fig. 5. The relationship between Chinese sentence intelligibility scores and STI and best-fit curves, the data of adults came from the study by Peng *et al.* (2011).

Table 2. The correlation coefficients and SDs for the scores versus STI.

Classroom	Grade 3		Grade 5	
	R	SD	R	SD
Word	0.94	3.5%	0.93	2.5%
Sentence	0.98	2.8%	0.98	2.4%

also plotted in Figs. 5 and 6. The data of adults came from the studies by PENG (2008) and PENG *et al.* (2011), respectively. Figures 5 and 6 show that both Chinese word identification and sentence intelligibility scores increase as STI and age increase.
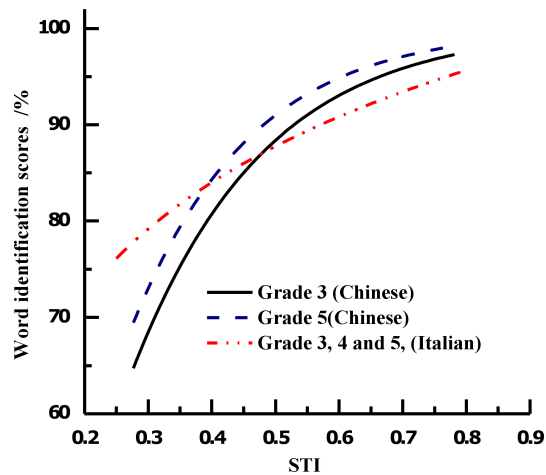

Fig. 6. Comparison of the curve between word identification scores and STI with Chinese and Italian for children.

4. Discussions

The results shown in Fig. 2 revealed that the older the children are the higher the word identification and sentence intelligibility scores under the same STI condition. The word identification in a sentence is constrained by phonetic input and context. In general, listeners can utilize different sources of knowledge about the structure of spoken language, including phonological, prosodic, lexical, semantic, syntactic, and situational, to constrain word choice. However, the age and language background (literacy) of children could affect children's speech identification in a room (NEUMAN, 1983). Sentence intelligibility test requires children to extract both word and semantic information from the test speech signal, as well as acoustical information. In the present study, the age of grade 3 and grade 5 students was from 8 to 11 years old, however, their auditory function, which gradually develops with the age, may not mature until 13 to 15 years old (NEUMAN, 1983). Younger children are less able to use stored phonological knowledge to reconstruct degraded speech input by reverberation and noise than older children and adults. Additionally, younger children are particularly vulnerable to perceptual difficulties in the presence of acoustical interference such as long reverberation and high background noise level. They therefore require better acoustical environments than adults do to achieve equivalent word identification and sentence intelligibility scores. Besides, younger children find it harder to concentrate on the beneficial auditory information (TALARICO *et al.*, 2006), which could

lead to the differences of speech identification ability among the different age groups. The combination of the factors mentioned above brought about that Chinese word identification and sentence intelligibility scores increased with the increase of age. It additionally illustrates that the sentence intelligibility scores were lower than the word identification scores in 0 dBA SNR condition for grade 3 students.

Word identification scores for children were examined as a function of STI by ASTOLFI *et al.* (2012). They used a diagnostic rhyme test (DRT) to obtain the Italian word identification scores in existing classrooms with RT from 0.37 s to 1.54 s. In the study of ASTOLFI *et al.* (2012) grade 3, 4, and 5 students (nominally 8–10 years old) were considered as one group. The best-fit curve between STI and Italian word identifications scores was described by a logarithmic curve. The curve between STI and Italian word identifications scores from the group was compared with the results from grade 3 (8 to 9 years old) and grade 5 (10 to 11 years old) students in the present study (Fig. 6). It can be seen from Fig. 6 that all curves indicate speech intelligibility scores increasing with the increase of STI value. There is a critical value where Chinese and Italian word identification scores were the same. Chinese word identification tended to show higher scores than Italian word identification above the critical value. However, below the critical value, Italian word identification scores were higher than Chinese word identification scores. The STI critical value for the transition varied with age of students. For instance, for grade 3 students the critical value of STI was 0.39, for grade 5 students it was 0.48. The differences in word identification scores between Chinese and Italian are related to subjects, languages, type of noise source, test methods, fitting methods, word lists, speech level and test conditions.

In Fig. 5 the performance of the three age groups can be compared by considering the required STI for each group to achieve near-ideal conditions for speech communication. For the results of sentence intelligibility tests, a sentence intelligibility score of 99% correct is used to indicate near-ideal conditions, because 99% correct scores are readily achievable in high STI (high SNR and short RT) conditions. Figures 2b and 5 show that all three age groups are capable of obtaining higher scores than a 99% in very high STI conditions. The mean trend in Fig. 5 shows that the adults could achieve 99% correct sentence intelligibility scores for a STI of 0.51. However, the grade 5 students required 0.61 STI and the grade 3 students required 0.71 STI to obtain a mean sentence intelligibility score of 99% correct. The grade 3 and 5 students would need a 0.2 and 0.1 greater STI value to obtain the same sentence intelligibility scores as the adults. For word tests, BRADLEY and SATO (2008) pointed that a word identification score of 95% correct is used to indicate

near-ideal conditions. According to Fig. 4, to achieve 95% Chinese word identification scores, the STIs required for grade 3, grade 5 students, and adults are 0.66, 0.61, and 0.56, respectively. STI values for near-ideal conditions determined from sentence intelligibility score and word identification scores for grade 5 students are approximately the same in classrooms. However, there is a difference in the required STI values between word and sentence tests for grade 3 students and adults. For adults, the STI determined from the sentence intelligibility test was lower than from the word identification test. But for grade 3 students the STI determined from the sentence intelligibility test was higher than from the word identification test. The result seems to show that the required objective acoustical index determined by a certain threshold of word identification scores might be underestimated for younger children (grade 3 students) in the classroom, but overestimated for adults. In fact, it can be found from Fig. 4 that a STI of 0.71 and 0.51 corresponding Chinese word identification scores were about 96% and 93% for grade 3 students, and adults, respectively. The result demonstrates that a 93% word identification score can achieve a 99% the sentence intelligibility score for the adults, but for grade 3 students, a 96% word identification score was needed to obtain a 99% sentence intelligibility score. This shows that different age groups require different word identification scores to obtain the same sentence intelligibility scores. Younger children (grade 3 students) need a higher word identification score to achieve near-ideal conditions (99% sentence intelligibility score), comparable to adults. Moreover, in the classroom it is more critical for children to clearly identify a sentence and to understand its meaning, than to recognize the single word. The method that required objective acoustical index (such as STI and SNR(A)) for different age group were determined by a 95% word identification score (Bradley, Sato, 2008; Astolfi *et al.*, 2012; Peng *et al.*, 2015), might not be optimal for children in classrooms. A certain threshold based on the sentence intelligibility score for different age groups to determine the required objective acoustical index may be more objective and practicable. This showed that a method based on the sentence test is more useful for speech communication evaluation in classrooms than that based on the word test for different age groups. To achieve the optimum sentence intelligibility and the best speech communication in classroom, the younger children require a higher STI than older children and adults.

5. Conclusions

The word identification and sentence intelligibility were evaluated by grade 3 and grade 5 students in three realistic primary classrooms with different RTs under different SNRs conditions. The word identification and sentence intelligibility scores were obtained and analyzed under different test conditions. The results showed that both Chinese word identification and sentence intelligibility scores for grades 3 and 5 children in the classroom increase with the increase of SNR, decrease with the increase of RT, and also increase with the increase of the age of students. Compared with Italian, Chinese word identification for grade 3 and grade 5 students tended to show higher scores than Italian word identification for the group with grade 3, 4 and 5 students above the critical value. But, below the critical value, Italian word identification scores for the group with grade 3, 4 and 5 students were higher than Chinese word identification scores for grade 3 and grade 5 students. The STI critical value for grade 3 and grade 5 students was 0.39 and 0.48, respectively. The relationship between STI and Chinese word identification and sentence intelligibility scores showed that the scores increase as the STI and the age increase. To achieve 99% sentence intelligibility score, the STIs required for grade 3, grade 5 students, and adults are 0.71, 0.61, and 0.51, respectively. The required objective acoustical index determined by a certain threshold of word identification scores might be underestimated for younger children (grade 3 students) in classroom, but overestimated for adults. A method based on the sentence test is more useful for speech intelligibility evaluation in the classrooms than that based on the word test for different age groups. To achieve the optimum sentence intelligibility and the best speech communication in classrooms, younger children need more favorable classroom acoustical environment with a higher STI than older children and adults.

References

1. Anderson B.W., Kalb J.T. (1987), *English verification of STI method for estimating speech intelligibility of a communications channel*, J. Acoust. Soc. Am, **81**, 1982–1985.

2. Astolfi A., Bottalico P., Barbato G. (2012), *Subjective and objective speech intelligibility investigations in primary school classrooms*, J. Acoust. Soc. Am., **131**, 247–257.

3. Brachmański S. (2004), *Estimation of logatom intelligibility with the STI method for polish speech transmitted via communication channels*, Archives of Acoustics, **29**, 4, 555–562.

4. BRACHMANSKI S. (2008), *Objective measure for assessment of speech quality in rooms*, Archives of Acoustics, **33**, 4(S), 177–182.

5. BRACHMANSKI S. (2012), *Automation of subjective measurements of logatom intelligibility in classrooms*, [in:] Automation, Florian Kongoli [Ed.], Rijeka, In-Tech, pp. 61–78.

6. BRADLEY J.S. (1986), *Speech intelligibility studies in classrooms*, J. Acoust. Soc. Am., **80**, 846–856.

7. BRADLEY J.S., SATO H. (2008), *The intelligibility of speech in elementary school classrooms*, J. Acoust. Soc. Am., **123**, 2078–2089.

8. GB 4959-85: *Methods of measurement for the characteristics of sound reinforcement in auditoria*, Standard of P. R. China, 1985.

9. HOUTGAST T. (1981), *The effect of ambient noise on speech intelligibility in classrooms*, Appl. Acoust., **14**, 15–25.

10. HYGGE S. (2003), *Classroom experiments on the effects of different noise sources and sound levels on long-term recall and recognition in children*, Applied Cognitive Psychology, **17**, 895–914.

11. KALIKOW D.N., STEVENS K.N., ELLIOTT L.L. (1977), *Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability*, J. Acoust. Soc. Am., **61**, 1337–51.

12. KITAPCI K., GALBRUN L. (2014), *Comparison of speech intelligibility between English, Polish, Arabic and Mandarin*, Proceedings of Forum Acusticum 2014, Kraków, available: http://www.fa2014.agh.edu.pl/fa2014_cd/article/RS/R20F_2.pdf.

13. KOCIŃSKI J., OZIMEK E. (2015), *Speech intelligibility in rooms with and without an induction loop for hearing aid users*, Archives of Acoustics, **40**, 1, 51–58.

14. NEUMAN A.C., HOCHBERG I. (1983), *Children's perception of speech in reverberation*, J. Acoust. Soc. Am., **73**, 2145–2149.

15. OZIMEK E., KOCIŃSKI J., KUTZNER D., SĘK A., WICHER A. (2013), *Speech intelligibility for different spatial configurations of target speech and competing noise source in a horizontal and median plane*, Speech Communication, **55**, 1021–1032.

16. PENG J.X. (2008), *Relationship between Chinese speech intelligibility and speech transmission index in rooms using dichotic listening*, Chinese Science Bulletin, **53**, 2748–2752.

17. PENG J.X., LÜ Y.X., LIU X.J., LU X.D. (2011), *Comparison of speech intelligibility between standard Chinese and English by native Chinese listener in rooms*, Journal of South China University of Technology (Natural Science), **39**, 130–134.

18. PENG J., YAN N., WANG D. (2015), *Chinese speech intelligibility and its relationship with the speech transmission index for children in elementary school classrooms*, J. Acoust. Soc. Am., **137**, 85–93.

19. STUART A. (2008), *Reception Thresholds for Sentences in Quiet, Continuous Noise, and Interrupted Noise in School-Age Children*, J. Am. Acad. Audiol., **19**, 135–146.

20. TALARICO M., ABDILLA G., ALIFERIS M., BALAZIC I., GIAPRAKIS I., STEFANAKIS T., FOENANDER K., GRAYDEN D.B., PAOLINI A.G. (2006), *Effect of age and cognition on childhood speech in noise perception abilities*, Audiology and Neurotology, **12**, 13–19.

21. TANG S.K., YEUNG M.H. (2004), *Speech transmission index or rapid speech transmission index for classrooms? A designer's point of view*, Journal of Sound and Vibration, **276**, 431–439.

22. TILLERY K.H., BROWN C.A., BACON S.P. (2012), *Comparing the effects of reverberation and of noise on speech recognition in simulated electric-acoustic listening*, J. Acoust. Soc. Am., **131**, 416–423.

23. WONG L.N., SOLI S.D., LIU S., HAN N., HUANG M.W. (2007), *Development of the Mandarin Hearing in Noise Test (MHINT)*, Ear & Hearing, **28(s2)**, 70–74.

24. YACULLO W.S., HAWKINS D.B. (1987), *Speech recognition in noise and reverberation by school-age children*, Audiology, **26**, 235–246.