# Impact of Languages to Speech Privacy and Intelligibility of Closed Spaces

**Y. Ma[a], D.J. Caswell[a], L. Dai[b] and J. T. Goodchild[c]**
a) Department of Mechanical and Manufacturing Engineering University of Calgary, Calgary, Alberta, Canada T2N 1N4
b) Industrial Systems Engineering, University of Regina, Regina, Saskatchewan, Canada S4S 0A2
c) SMED International, Calgary, AB T2C 4T5, Canada

## ABSTRACT

The present study investigates the speech privacy and intelligibility of closed spaces in multicultural environments. Most assessments for speech privacy and speech intelligibility among the current research rely on the subjective measurements utilized with the test materials of English and other Western languages. Effects of different languages and accents in speech privacy and speech intelligibility are usually overseen. Subjective measurements are conducted in this study for closed spaces by using English and a tonal language. The differences in speech privacy between the two languages are evident and significant. It is also found in this study that the existing single word tests used in research and industrial practice for subjectively evaluating speech privacy need modification when closed spaces are considered. The subjective measurement results of this study are compared with the objective measurement index AI.

## RÉSUMÉ

Le but de cette recherche est d'étudier l'intimité et l'intelligibilité des conversations dans des espaces clos en milieu multiculturel. La plupart des études se basent sur des données subjectives utilisées dans l'étude des tests basés sur la langue Anglaise et d'autres langues Occidentales. L'influence des accents et d'autes langues sur l'intimité des conversations et l' intelligibilité des conversations ne sont géneralement pas prises en compte. Des experiences subjectives ont été conduites dans cette étude dans des espaces clos en utilisant L'Anglais et une langue tonale. Les differences dans le cadre de l'intimité des conversations entre les deux languages sont évidentes et significatives. Les résultats montrent aussi que le test basé sur la pronunciation d'un seul mot tel qu' utlisé en recherche et dans le milieu industriel pour l'évaluation subjective, doit être modifié lorsque des espaces clos sont considérés. Les résultats de cette étude sont comparés avec les mèthodes objectives utilisées dans l'indice d'articulation (IA).

## 1. INTRODUCTION

Speech privacy is the opposite concept of speech intelligibility. The lower the speech intelligibility is, the higher is the speech privacy. With this consideration, speech privacy can be assessed by the predictors of speech intelligibility. Generally, there are two families of methods for characterizing the speech intelligibility: the subjective or direct method and the objective or indirect method. The subjective methods involve human subjects in a procedure known as speech intelligibility testing. Listeners are placed in real listening conditions and must transcribe each proposed language unit as they perceive it. The intelligibility score is then derived from the count of the correctly transcribed units. These direct methods are problematic in their reliability and their reproducibility.

The objective methods generally yield an intelligibility index from a measurement of physical characterization of acoustical environment. With these methods, each index is distinguished from the other indices according to the measurement it is based on. There are three types of acoustical measures widely used in the research to date [1-4]: articulation index (AI) or speech intelligibility index (SII) [5-8], speech transmission index (STI) [9,10] and sound early-to-late ratio ($C_{50}$) [11]. Among the recent research, investigations are also found on intelligibility of rooms [12-15].

It should be noticed that most objective assessments for speech privacy and speech intelligibility are evaluated by the subjective measurements. The subjective measurements, however, are mainly based on studies of English and other Western languages as per the existing standards adopted in research and industrial practice for evaluating speech privacy. The impact on speech privacy by other languages and accents is overseen. In the modern society, however, environments involving different languages and accents are common especially in the places such as international organizations, government and business offices, classrooms and medical clinics of multicultural communities. A systematic investigation on the multilanguage and accent impact on speech privacy in closed spaces is therefore necessary. In this study, subjective measurements using both English and a tonal language, Mandarin are conducted. With the two languages, the differences in speech intelligibility, and hence speech privacy, are studied. The subjective measurement results are also compared with the objective measurement index AI obtained by ASTM E1130 method [6], which is widely used in evaluating speech intelligibility and speech privacy.

## 2. EXPERIMENTAL MEASUREMENTS

### 2.1 Objective measurements

The articulation index (AI) method based on the existing standard ASTM E1130 is employed in this study. Noise signals are captured with B&K Dual Channel Real-time Analyzer type 2133, microphones 4189 and 4133, signal amplifier, 1/3 octave filter and speaker. Figure 1 illustrates the experimental setup scheme. Pink noise is used as the sound source in the talker room B and the sound pressure levels in one-third octave bands are measured in the listener room A. In each listener room used in the experiments of this study, at least 4 listener positions are used for the measurements. The measurements and data process procedures are conducted according to ASTM E 1130.

For simulating the real environment of the closed spaces, the measurements are conducted in some typical office rooms of a university building. The volume of office rooms range from 37 – 56 $m^3$. The partition walls between the closed offices are typical drywalls from floor to ceiling; the ceilings are gypsum-ceiling boards with continuous plenum. The space furnishings in the rooms, where the measurements were conducted, are study tables, chairs, bookcases and file cabinets.

Background noise in the rooms is mainly due to computer-generated noise and rated from NC 37 to 41 in the closed offices. The reverberation time measured is in between 0.21 and 0.37 seconds in the listener rooms. The measurements of the background noise levels and reverberation time show that the occupied closed offices are under normal

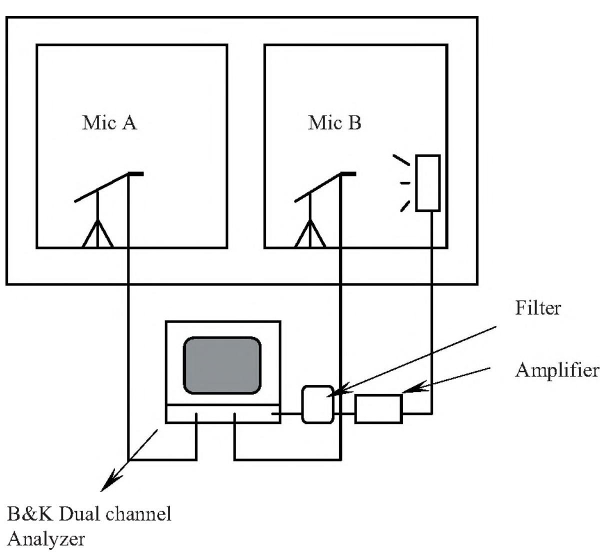


**Figure 1. Set-up for AI Measurement**

acoustical conditions and suitable for carrying out the speech intelligibility and speech privacy tests.

### 2.2 Subjective measurements

**a) Subjects and speech test materials**

All the measurements are conducted in typical closed office rooms and classrooms, and both the objective and subjective tests performed in the study are under the same background noise conditions. A total of 47 subjects take part in the speech intelligibility tests for English. They are all native English-speaking adults ranging in age from 19 to 25. A total of 22 subjects take part in the speech intelligibility tests using Mandarin. The subjects are all native Mandarin-speaking adults ranging in age from 25 to 40. All of the subjects showed no evidence of hearing problems.

In performing the experiments, three types of speech test materials are used for measuring speech intelligibility of English: (a) Modified Rhyme Test (MRT) words, (b) multi-choice conversations and articles, and (c) open-set sentence. MRT words are selected from standard ANSI S3.2-1989 [16], and consist of 50 six-word lists of monosyllabic English words. The large majority of the words used have three syllables in a consonant-vowel-consonant sequence.

Almost all the research to date use different test materials of single English words (MRT, PB word, and DRT etc.) in evaluating speech intelligibility. However, in real communication situations, people can understand the whole sentence or conversation even if they just hear a few words. This is due to the coherence of human being in the context, speech tone and intonation when people are speaking. Therefore, the understanding of a whole sentence appears to be more representative for a realistic communication situation than the intelligibility of discrete words that have no logical relation among them. To simulate the communication of human society, new test materials should be utilized in the experiments to represent the real world communication environments. In this study, multi-choice conversations and articles selected from standard English listening tests are employed. The test materials employed consist of short and long conversations together with articles. In the experiments utilizing the conversations and articles, the listeners are required to listen to short and long conversations between two people or a vocal lecture article. After each conversation or reading of an article, there is one or several spoken questions about the listening materials and four answer choices corresponding to each of the questions. All the listening materials are related to common people's communication happening in everyday life. In each test, there are 50 questions to be answered. The open-response-set sentences are also employed in the experiments. In this type of experiments, the listeners are required to either repeat or write down what is heard. For example, on the recording the listener will hear:

"*He's sick of his job*".

In the question sheet, the listener will read:

(A) He doesn't like his work;

(B) He isn't doing a good job;
(C) He became ill at the office;
(D) He's tired of looking for a job.

If the listener hears and comprehends the sentence, he or she may learn from the sentence that the man doesn't like his job, and thus select the best answer (A). In performing the tests, each experiment used includes 30 such open-response-set questions.

Similar to the English test materials, the single words, articles, real conversations and open-set sentences are used in speech intelligibility and privacy experiments of Mandarin. Most English words are multi-syllable. A single syllable of an English word has relatively few phonemes and is usually no meaning. In contrast to English, one Mandarin character has one syllable associated with it and a single character is mostly meaningful and tonal [17]. However, in many cases, one syllable of Mandarin may represent different characters. Tones are extremely important in Mandarin speech because the different tonality of the same monosyllable will give different meaning. There are four tonal patterns in Mandarin: tone 1 (flat tone), tone 2 (rising tone), tone 3 (falling-rising tone) and tone 4 (falling tone). In single word test, there are 40 five-word lists are used; each list includes five single words (characters) which have the same tone and same initial consonant but different vowels. In the experiments with conversations or article materials, the standard Mandarin listening test in Chinese Proficiency Test (HSK) is used. After each conversation or article reading, there will be one or several spoken questions about the listening materials and there are four answer choices for each of the questions. In each experiment, there are 40 questions to be answered. The open-set test consists of several conversations or short articles, the listeners are asked to answer what they heard (no multi-choices are provided). There are total 15 questions in an open-set test.

For both languages, the multi-choice conversations, articles and open-set sentences used in the test are very similar to those listening comprehension tests, such as TOEFL and HSK, which are widely accepted by the academic institutions for evaluating listening comprehensions. The selection of the test materials is based on the considerations that the materials may best reflect realistic communication situations and people's subjective judgement in communication.

**b) Subjective measurement procedures**

Both English and Mandarin subjective measurements are conducted in the same office pairs as the objective measurements as illustrated in Figure 1. In the closed office environments, one or two talkers sit in the middle of the talker's room, and at least four listeners sit at different positions in the listener's room. Firstly, one of the talkers reads the test materials or two talkers have the conversations in "normal voice level", which corresponds to the voice level used in normal conversation and has an overall level of about 58 dBA. The talkers then read in "raised voice level", which is the voice level often used in addressing the people in a regular classroom or speaking into a microphone about one meter away. In this case, the voice has an overall sound level of about 64 dBA. A recording tape is also played at about 58 and 64 dBA respectively in the talker room. In both situations, the listeners are asked to answer the questions of the test materials on the response sheets. Furthermore, for both languages, the talkers are asked to speak at a speed as that they use in a normal conversation, similar to the speed used for the listening comprehension tests. The test results did not show obvious differences between male and female talkers. The aim of the present research is to assess the language impact to speech privacy in a communication environment close to that of real world including both males and females. However, the gender differences is planned for further studies in future investigations.

It is noted in the tests, that Mandarin shows higher long-term overall sound levels than other languages, in general. This agrees with the results reported by Byrne *et al* [19] in which a comparison of twelve languages (including American English) was performed. Mandarin was found to have the highest average level of 75.2 dB at 20 cm from a speaker's mouth, in comparing with an average of 72 dB of the twelve other languages used for the tests.

## 3. RESULTS AND DISCUSSION

**a) Objective and subjective measurements**

The final speech intelligibility scores of the listeners are collected and computed as the arithmetic average of the percent of the correct answers identified by the listeners. The final scores are then accepted as a score representative of the specific listener room.

Table 1 lists the AI results of the objective measurements in four pairs of closed offices. The results of the subjective measurements as the percent of intelligibility for different test materials are listed in Table 2.

According to the standard ASTM E 1374 "Standard Guide for Open Office Acoustics and Applicable ASTM Standard" and ASTM E 1130 "Standard Test Method for Objective Measurement of Speech Privacy in Open Offices Using Articulation Index", together with the classical relationship between AI and speech intelligibility, different level of speech privacy can be conveniently categorized as follows [6]:

1. Confidential Privacy: AI $\leq 0.05$. Speech can be detected but not understood by the receiver. It also implies less than 10% word and 5% sentence intelligibility.
2. Normal Privacy: $0.05 < AI \leq 0.20$. Effort is required for the receiver to understand the speech.
3. Transitional Privacy: $0.20 < AI \leq 0.40$. The corresponding speech is mostly understood and can be distracting.
4. No Privacy: AI $> 0.40$. Speech is clearly understood by the receiver.

From Table1, the objective measurement results show

**Table 1. Results of Subjective Measurements and AI Values**

| Office | No. Sub. | AI (Avg) | Privacy Type |
|---|---|---|---|
| Closed 1 | 16(E) | 0.001(N) | Conf. |
| | 4 (C) | 0.04(R) | Conf. |
| Closed 2 | 5 (E) | 0.001(N) | Conf. |
| | 4 (C) | 0.01(R) | Conf. |
| Closed 3 | 9 (E) | 0.001(N) | Conf. |
| | 5 (C) | 0.01(R) | Conf. |
| Closed 4 | 7 (E) | 0.002(N) | Conf. |
| | 6 (C) | 0.01(R) | Conf. |

Note: N indicates the normal voice level and R the raised voice level as indicated above. E indicates result for English and C for Mandarin.

**Table 2. Results of Subjective Measurements and Percent of Intelligibility**

| Office | Percent of Intelligibility | | |
|---|---|---|---|
| | **Single Word** | **Multiople-choice Conversation** | **Open-set Sentence** |
| Closed 1 | 28%(EN)<br>40%(CN)<br>49%(ER)<br>67%(CR) | 2%(EN) 1%(CN)<br>45%(ER) 47%(CR) | 35%(ER)<br>44%(CR) |
| Closed 2 | 39%(EN)<br>44%(CN)<br>53%(ER)<br>71%(CN) | 1%(EN) 2%(CN)<br>28%(ER) 38%(CR) | |
| Closed 3 | 36%(EN)<br>42%(CN)<br>43%(ER)<br>64%(CR) | 1%(EN) 2%(CN)<br>12%(ER) 47%(CR) | |
| Closed 4 | | 25%(ER) 40%(CR) | 28%(ER)<br>41%(CR) |

that the AI values are all less than 0.05 at normal or raised voice levels in four test offices. This implies that the speech privacy is Confidential Privacy as categorized above. However, from the comparison of AI value and privacy type with speech intelligibility scores of both English and Mandarin as exhibited in Figure 2, it can be seen that there is a good correspondence between the privacy type and the conversation test results at normal voice level. The values of percent of intelligibility for both the English experiments and Mandarin ones are close and the values are varying with the same trends for difference test materials. But this correspondence does not exist in the tests with the raised voice level.

For single word experiments, the Percent of Intelligibility is very high even at normal voice levels as shown in Figure 2. However, the values of Percent of Intelligibility obtained from the experiments with conversations or articles are significantly lower than that of single word experiments. It should be noted that all the Percent of Intelligibility values of both the single word and conversation/article experiments are measured under identical experimental conditions. This may raise fundamental concerns in assessing speech privacy and intelligibility of closes spaces. Speech intelligibility measures the degree of comprehension in a common verbal communication of real world. In the communications of reality, one may understand the meaning of a sentence or a conversation by hearing and comprehending a few "key" words embedded in the conversation. On the other hand, one may completely lose the meaning of a sentence or conversation if he or she only catches some of the words in the sentences and misses the "key" words, even though that the words caught are fully understood by the listener. The

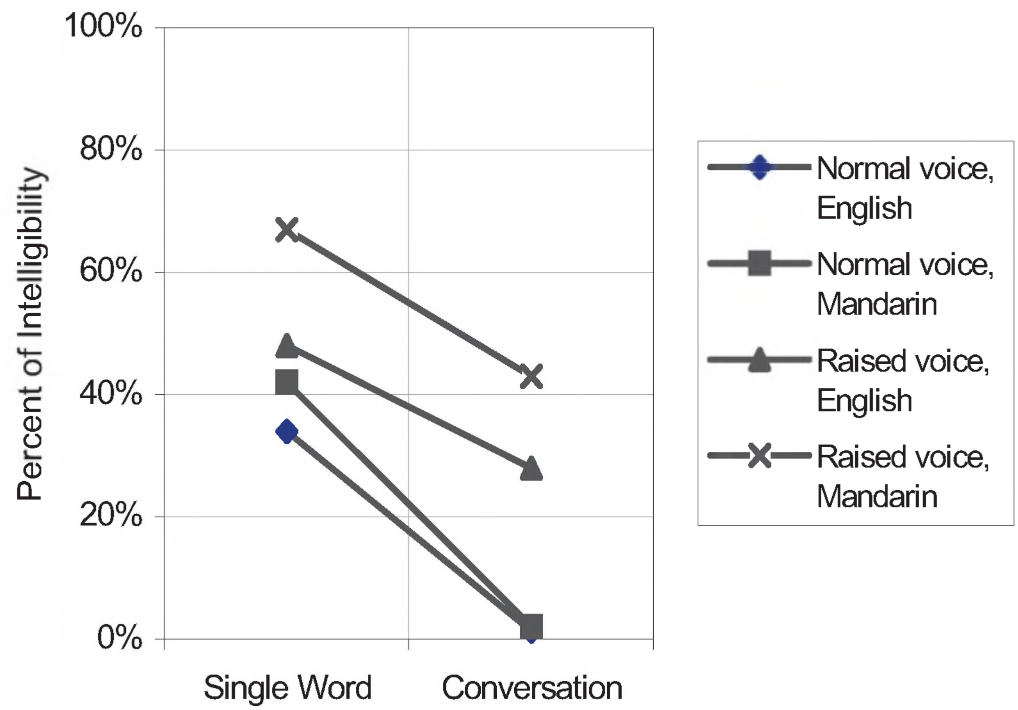


**Figure 2. Speech Intelligibility Scores of Test Materials**

reliability of speech intelligibility evaluated by single words may get even worse in considering that some words, especially for many of English words, are single-syllabled with few phonemes and many of them are sibilant. It appears that the measurements for speech privacy or intelligibility with single words may not reflect the actual level of speech privacy or intelligibility in reality, and not very reliable for subjectively measuring the speech privacy in the environments of closed spaces. Modifications on the testing materials to be used in the speech privacy measurements seem necessary.

Voice effort level also affects speech intelligibility and privacy significantly in both languages. With respect to English, the percent of intelligibility is increased by 14% for single words, and 27% for conversations comparing with

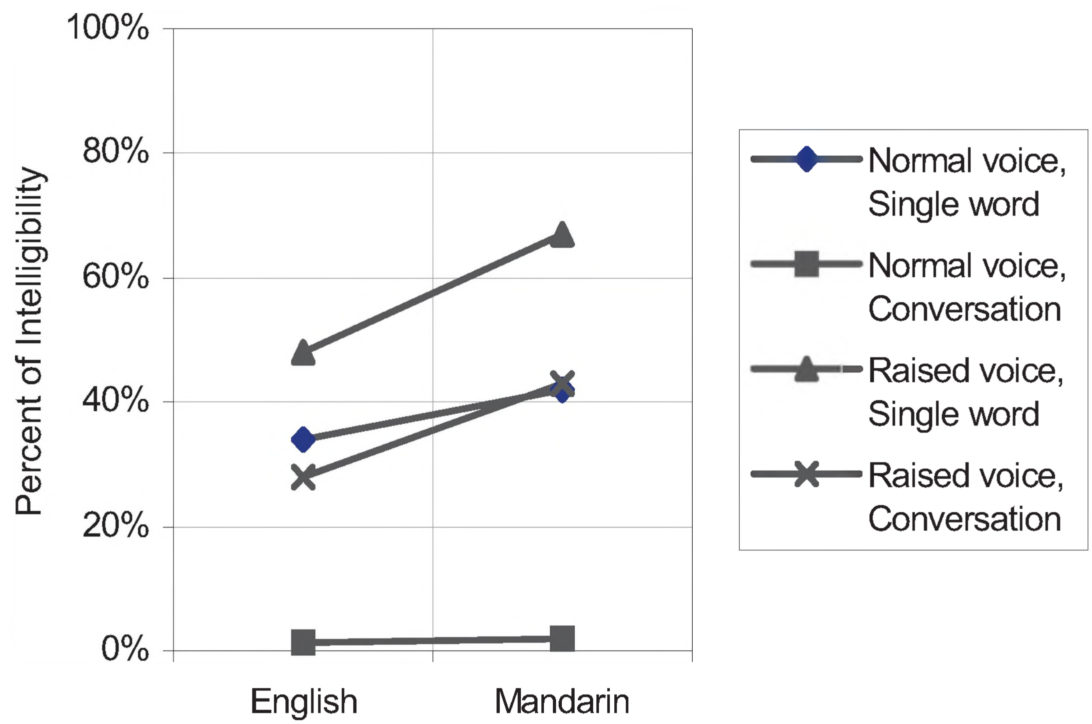


**Figure 3. Speech Intelligibility Scores of English and Mandarin**

the raised voice level with normal voice level; for Mandarin, the corresponding scores is increased about 22% for single words and 41% for conversations. It may also be observed from Table 1, Table 2 and Figure 2, at the raised voice level, a substantial difference exists between the subjective test and the AI measures. It should be noted that the calculation procedures for obtaining the standard AI values are originally established under open space environments, and the weight factors required for determining AI corresponding each band frequency are also determined under the open space environments. The results of this study suggest that the modification is needed to the standard methods in determining AI values of a closed office.

In the computing process for determining AI at raised voice level, it was found that the biggest contribution to the AI value came from low frequencies between 250 to 1000Hz, which are mainly from the vowel sounds in both languages. However, one may keep in mind that the closed partition between sound source and receiver has the direct effects to the sound transmission characteristics as well as the final contribution of each frequency band to the AI results.

It is found in the experiments that the spectrum of voice heard on the other side of the closed partition wall is modified by filtering effect of transmission loss and by some of the resonance effects of the partition wall. Further investigation needs to be carried out on the applicability of the standard AI method for assessing speech privacy in closed office at raised voice level.

**b) Comparison of English and Mandarin**

Figure 3 shows the average subjective measurement results in four closed offices for English and Mandarin. It can be seen from the figure that the single word intelligibility of Mandarin is generally better than that of English at normal and raised voice levels. It is recognized in the experiments that some English consonants, such as fricatives, nasals, which are presented in high frequencies and provide significant information, are blocked by the closed partition wall in closed spaces. As reported in [18], the tones of Mandarin help to recognize vowels and consequently increase the word intelligibility. The same situation happens in conversation intelligibility and open-set sentence intelligibility experiments at raised voice. The scores of intelligibility of conversation at normal voice are very close for both languages. This is because that the long conversations mask the effect of Mandarin tone recognition as that for the individual word. The intelligibility of whole conversation therefore presents the similar results as that of English. In general, the results of the experiments show that Mandarin has better intelligibility or worse speech privacy than English in closed space environments.

## 4. CONCLUSIONS

Language impacts on speech privacy and intelligibility in closed spaces such as offices were investigated in the present study. Based on the subjective and objective experiments performed and the analyses carried out in the study, the following can be stated.

1. A series of subjective experiments are conducted in closed spaces with employment of English and a tonal language, Mandarin. The objective measurements with the Articulation Indices (AI) are also performed and compared with the subjective measurement results for speech intelligibility in closed spaces. Similar studies are not found in the current literature.
2. The experimental results suggest that single word tests seem not very suitable for subjective measurements of speech privacy evaluation in closed spaces, regardless of the languages used.
3. The results of ASTM standard method are consistent with subjective test only at normal speech voice level. However, differences are found at raised voice levels for both English and Mandarin.
4. The results of the study also showed the evaluation differences in intelligibility and speech privacy in closed spaces between the two different languages. Generally, as found in the experiments and analyses, Mandarin has better intelligibility or worse speech privacy than that of English in closed space environments. The language impact on speech privacy and intelligibility in a closed environment is evident. Hence, speech privacy is actually language dependent. Strictly speaking, different language has different speech indelibility rate under the identical environment. The language impact must therefore be considered in future office design.
5. Several interesting acoustic characteristics are found in this study that is significant in the acoustical designs for closed offices to be used in multicultural environments.

## REFERENCES


1. French, N. R. and Steinberg, J. L., 1947, "Factors governing the intelligibility of speech sounds", J. Acoust. Soc. Am., **19**, 90-119.
2. Bradley, J. S., 1986, "Auditorium Acoustics Measures From Pistol Shots", J. Acoust. Soc. Am., **80**, 199-205.
3. Bradley, J. S., 1986, "Predictors of Speech Intelligibility in Rooms", J. Acoust. Soc. Am., **80**, 837-845.
4. Bradley, J. S., Reich, R. D. and Norcross, S. G., 1999, "On the Combined Effects of Signal-To-Noise Ratio and Room Acoustics on Speech Intelligibility", J. Acoust. Soc. Am., **106**, 1820-1828.
5. ANSI (1997). ANSI S3.5-1997, "Methods for calculation of the speech intelligibility index", American National Standard Institute, New York (1997).
6. ASTM (1994). E1130-90 (R1994), "Objective measurement of speech privacy in open office using articulation index", American Society for Testing and Materials, Philadelphia (1994).

7. Kryter, K. D., 1962a, "Methods for the calculation and use of the articulation index", J. Acoust. Soc. Am., **34**, 1689-1697.
8. Kryter, K. D., 1962b, "Validation of the articulation index", J. Acoust. Soc. Am., **34**, 1698-1702.
9. Houtgast, T., Steeneken, H. J. M. and Plomp, R., 1980, "Prediction speech intelligibility in rooms from the modulation transfer function: I General room acoustics", Acustica, **46**, 60-71.
10. Houtgast, T. and Steeneken, H. J. M., 1985, "A review of the MTF concept in room acoustics and its use of estimating speech intelligibility in auditoria", J. Acoust. Soc. Am., **77**, 1069-1077.
11. Cavanaugh, W. J., Farrell, W.R. and Hirtle, P.W., 1962, "Speech privacy in buildings", J. Acoust. Soc. Am., **34**, 475-492.
12. Tisseyre, A., Moulinier, A. and Rouard, Y., 1998, "Intelligibility in various rooms: comparing its assessment by (RA)STI measurement with a direct measurement procedure", Applied Acoustics, **53**, 179-191.
13. Warnock, A. C. C., 1973, "Acoustical privacy in the landscaped office", J. Acoust. Soc. Am., **53**, 1535-1543.
14. Wyerman, B. R., 1982, "A study of speech privacy in closed offices", J. Acoust. Soc. Am., **72**, 436-442.
15. Dai, L., Ma, Y., and Caswell, D., "Experimental Assessment of Speech Privacy and Intelligibility in Multi-Language Environments," ASME International Mechanical Engineering Congress and Exposition, New Orleans, Louisiana, 2002.
16. ANSI (1995). ANSI S3.2-1989, "Method for Measuring the Intelligibility of Speech over Communication Systems", American National Standard Institute, New York (R1995).
17. Fu, Q.J., Zeng, F.G., Shannon, R.V., and Soli, S.D., 1998, "Importance of tonal envelope cues in Chinese speech recognition", J. Acoust. Soc. Am., **104**, 505-510.
18. Kang, Jian, 1998, "Comparison of speech intelligibility between English and Chinese", J. Acoust. Soc. Am., **103**, 1213-1216.
19. Byrne, D., Dillon, H., Tran, K., Arlinger, S., Bamford, J. Wilbraham, K., Cox, R., Alexander, G., Hageman, B., Hetu, R., Kie, J., Lui, C., Kiessling, J., Kotby, N., Nasser, N., El Kholy, W., Nakanishi, Y., Oyer, H., Lambert, C., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G. (Moscow), Westerman, S., Ludvigsen, C., 1994, An international comparison of long-term average speech spectra, J. Acoust. Soc. Am., **96**, 2108-2120.