

# Logatome and Sentence Recognition Related to Acoustic Parameters of Enclosures

Jędrzej KOCIŃSKI, Edward OZIMEK

*Institute of Acoustics, Faculty of Physics  
Adam Mickiewicz University*

Umultowska 85, 61-614 Poznań, Poland; e-mail: jedrzej.kocinski@amu.edu.pl

(received January 13, 2017; accepted March 27, 2017)

The paper deals with relationship between speech recognition and objective parameters of enclosures. Six enclosures were chosen: a church, an assembly hall of a music school, two courtrooms of different volumes, a typical auditorium and a university concert hall. Dirac 4.1 software was used to record impulse responses (IRs) in the chosen measurement points of each enclosure. On this base, the following objective parameters of the enclosure were determined: Reverberation Time (RT), Early Decay Time (EDT), Weighted Clarity ( $C_{50}$ ) and Speech Transmission Index (STI). A convolution of the IRs with logatome tests and the Polish Sentence Test (PST) was made. Logatome recognition and speech reception threshold (SRT – i.e., SNR yielding 50% speech recognition) were evaluated and their dependence on the objective parameters were determined. Generally a linear relationship between logatome or SRT and RT or EDT was found. However, speech recognition was nonlinearly related (according to psychometric function) to STI values. The most sensitive range of the logatome and sentence recognition relative to STI changes corresponded to the middle range of STI values. Below and above this range, logatome and sentence recognition were much less dependent of STI changes.

**Keywords:** STI; speech intelligibility; reverberation time.

## 1. Introduction

Relationship between speech recognition and the acoustic parameters of an enclosure has been studied by many authors (BRACHMAŃSKI, 2004; 2008; BRADLEY, 1986; BRADLEY *et al.*, 2003; HOUTGAST, STEENEKEN, 1985; HOUTGAST *et al.*, 1980; JACOB *et al.*, 1991; STEENEKEN, HOUTGAST, 1980; YANG, BRADLEY, 2009), however, it is still a challenging topic. Particularly often speech intelligibility under different reverberation conditions was studied. The obtained data are not fully in agreement and it is difficult to draw unequivocal conclusions. Though, the results generally indicate that word intelligibility in quiet decreases versus reverberation time (RT) and in the presence of masking noise decreases more rapidly. Further studies are needed to gain more detailed insight into the effect of reverberation and noise on speech intelligibility. The approach for solving some problems was introduced by BRACHMAŃSKI (2008).

An interesting measure in room acoustics is the Modulation Transfer Function (MTF) (HOUTGAST,

STEENEKEN, 1973), which represents a degree of changes in modulation depth of the temporal envelope of a signal propagating in a room as a function of modulation frequency. The MTF function, which essentially is the Fourier transform of the impulse response of a room, allows to derive a single measure – Speech Transmission Index (STI), which makes possible to characterize the effect of reverberation time and noise on speech listening quality which can be expressed in intelligibility score. However, interpretation of STI in terms of intelligibility score is a complex matter since it depends not only on the room parameters but also on other factors such as type of speech material used, listeners, subjects and the language. For example in Polish some research in this area can be found in (BRACHMAŃSKI, 2015; MAJEWSKI *et al.*, 1998).

Many different speech recognition tests such as: one-syllable words rhyme tests, words logatomes or numbers have been introduced in the last decades (BRACHMAŃSKI, STARONIEWICZ, 1999; HAGERMAN, 1982; HOUSE *et al.*, 1965; KALIKOW *et al.*, 1977; KOLLMEIER, WESSELKAMP, 1997; NILSSON *et al.*,

1994; PRUSZEWICZ *et al.*, 1994a; 1994b; VERSFELD *et al.*, 2000; BRACHMAŃSKI, 2004). Most of these tests, however, do not reflect a real communication process in which the basic units are sentences (OZIMEK *et al.*, 2006; 2009a; 2009b; PLOMP, MIMPEN, 1979). Measurement of the subjective quality of an enclosure in terms of speech recognition should be carried out using a most natural signals such as sentence tests. Therefore, a sentence test was used in the present study to measure speech recognition in a room. Interesting suggestion is taking into account a measure of recognition expressed as Speech Reception Threshold (SRT), defined as the SNR corresponding to 50% speech recognition. This measure is more phonemically representative for a given language and was proven to give more accurate speech recognition data than standard word tests. A new adaptive method based on speech masking is introduced here. The target (speech) signal was convoluted with a room impulse response (IR) measured in the enclosure using a dummy head and speech reception thresholds (SRTs) in noise were determined for such convoluted (reverberated) sentence material. The SRT measurements were recently used by GEORGE *et al.* (2010) in investigations of the effects of reverberation and noise on sentence recognition for hearing-impaired subjects. To compare the results of the proposed method with “classic” one, a logatome recognition was also determined. Next, both subjective measurements were compared with objective parameters gathered in all rooms.

Particular attention is paid to the Speech Transmission Index (STI), introduced by HOUTGAST and STEENEKEN (1973), which is some sort of measure describing the effects of reverberation and noise on speech intelligibility. Technically, STI is calculated as the weighted sum of Modulation Transfer Function (MTF) indices, one for each octave frequency band from 125 Hz through 8 kHz, where each MTF value is derived from MTF values over 14 different modulation frequencies, taking into account auditory effects according to IEC 60268-16 norm (2011).

## 2. Purpose

The basic purpose of the present study is to find a relationship of the logatome and sentence recognition versus following objective measures of rooms: RT, EDT,  $C_{50}$  and STI. Six enclosures as the representatives of different kinds of public buildings were chosen: a church, an assembly hall of a music school, two courtrooms of different volumes, a typical auditorium and a university hall, which is also used by the Philharmonic Orchestra. The obtained data generally showed that the logatome and sentence recognition tested in these enclosures were linearly (RT, EDT) or nonlinearly (STI) related to the objective parameters. The paper is organized as follows. First section describes

the method of gathering the objective and subjective parameters. Then, objective and subjective measures for tested enclosures are shortly described. Finally, the analyses of the results and the relationship of subjective and objective measures are presented.

## 3. Method

### 3.1. Recordings

A PC with B&K Dirac 4.1 software was used to record and collect IRs of the tested enclosures. This software additionally allowed to calculate the objective parameters of a room according to ISO 3382 norm (2010). To extract an IR, a Maximum Length Sequence (MLS) (BORISH, ANGELL, 1983; CHU, 1990; KUTTRUFF, 2009) technique was used as a driving signal instead of an impulse burst (5 averagings were used). During the RT, EDT and  $C_{50}$  measurements, the MLS signal was generated by the software and fed via a D/A converter (ESI U2A) to the amplifier and then to the omnidirectional sound source (IZT 10/12A) placed at different places of the enclosure at the height of 1.5 m. The measurements of the omnidirectional source made in an anechoic chamber proved its properties which are in line with ISO 3382 norm (2010). For the STI measurements Yamaha HS50M was used instead of omnidirectional source according to the IEC 60268-16 norm (2011) at the same height. For all the enclosures the most common places of sound source were chosen as well as many different positions of a receiver (9 in church, 6 in Adam Mickiewicz University Hall, 14 in large courtroom, 9 in auditorium, 18 in assembly hall, 6 in small courtroom). Two different types of receivers were used, namely an omnidirectional microphone (Svantek SV01A) and a dummy head (Neumann KU100). The former was used to get the objective parameters, while the latter was used to collect the IRs via a head (with a HRTF). According to ISO 3382 norm (2010) the receivers were placed at the height of 1.2 m which corresponds to the height of ears of a sitting person.

IRs convolved with the test material allowed to measure logatome and sentence recognition in the laboratory (LONGWORTH-REED *et al.*, 2008; PENG, 2008; YANG, 2006).

### 3.2. Objective parameters

Four objective parameters in each enclosure were measured: reverberation time (RT), according to ISO 3382 norm (2010), early decay time (EDT), weighted clarity  $C_{50}$  (MARSHALL, 1994), and speech transmission index (STI), introduced by HOUTGAST and STEENEKEN (1973).

Since the background levels were less than 40 dB(A) (the rooms were empty, thus there was no

additional noise and there was no air conditioning nor active ventilation system) in all rooms it was possible to measure T30 according to ISO 3382 norm (2010). Those values were taken as the RTs. RT gives the information on the diffuse sound decay and are derived from the decay curve between 0 dB and 60 dB below the initial level. EDT is derived from the decay curve section between 0 dB and 10 dB below the initial level. Since EDT relates more than the other reverberation parameters to the initial and highest level part of the decaying energy, it is also the parameter that relates most to modulation reduction, hence speech recognition reduction (Manual, 2008). Standard RTs and EDTs in six octave bands were measured and then the mean values were determined.

$C_{50}$  is the logarithmic early-to-late arriving sound energy ratio, where “early” means “during the first 50 ms” and “late” means “after the first 50 ms”. In this paper we use weighed  $C_{50}$  which was introduced by MARSHALL (1994). The weighting values reflect the impact of different octave bands on speech recognition and can be used for prediction of speech intelligibility according to the above mentioned paper: for the octave bands of 0.5 kHz, 1 kHz, 2 kHz and 4 kHz the weights of 0.15, 0.25, 0.35 and 0.25 were assigned respectively, and then the results were summed.

STI is based on the assumption that the distortion of amplitude modulation of the informative signal is crucial for speech intelligibility decrease. STI is the most comprehensive and important speech recognition parameter. It ranges from 0 (bad intelligibility) to 1 (excellent intelligibility).

### 3.3. Intelligibility tests

The logatome test (BRACHMANSKI, STARONIEWICZ, 1999) and the Polish Sentence Test (PST) (OZIMEK *et al.*, 2009a) were used to measure speech recognition in enclosures. The so-called auralization was used to present the tests, namely the IRs recorded via dummy head in all measurement points of all enclosures were convolved by a computer program with the elements of speech intelligibility tests (logatomes or sentences), and presented to the subject via Tucker-Davis Technology (TDT) RP2 (D/A converter) and Sennheiser HD580 headphones in the acoustically insulated booth. The listening sessions were controlled using Matlab 6.5 software. In the case of logatome test the total level of the target signal presentation was equal to the level measured in the enclosure during the recordings; thus, all the *in situ* conditions were preserved. For logatome recognition tests a commonplace determination of the ration correctly understood elements to all presented elements was determined. Each subject was listening to 50 logatomes in each reverberant condition and measurement point and was asked to write down the logatome he or she heard out. Af-

ter that next logatome was played back. After a set of 50 logatomes, for a particular condition the logatome recognition was calculated as a percentage of correct responses.

Apart from the logatome test, the Polish Sentence Test presented at the background of masking noise was also used. However, a different approach was taken here. Instead of percentage of correctly understood elements (like for logatomes) SRT was determined in masking conditions. To measure the SRT the sentences were masked by additional noise (so-called babble noise) which was artificially created and played back together with the reverbered sentence. By definition SRT is the SNR for which 50% of elements were correctly understood. Instead of using commonplace constant stimuli method, a standard 1-up/1-down adaptive procedure (LEVITT, 1971) was used to determine SRT. SNR was varied adaptively with respect to the most recent subject's response. The SNR was either increased or decreased by some value (step) when the most recent response was incorrect (1-up) or correct (1-down), respectively. This method, by definition, converges the SNR value to the 50%-equilibrium point on the recognition function, i.e. the SRT. To make the method faster, a 2-dB step was used until the first incorrect answer was recorded; then it was changed to 1dB. The so-called babble noise, made from the mixture of all sentences used in the test, was taken as a masker (see OZIMEK *et al.*, 2009c for details). The power spectrum of the babble noise optimally matched the power spectra of the sentences. SRT was calculated as the mean of the last 8 (from 13) nominal SNRs. Each sentence was convolved with the IR from a particular measurement point and then presented to a subject together with babble noise. The subject's task was to listen to the sentence. After each sentence a panel with 18 words was shown on the screen of the computer and the subject was asked to click on the words he or she had just heard to compose a sentence. This method was introduced by OZIMEK *et al.* (2009c) and was proven to give the same results as the standard adaptive method, however it is automated, thus faster and the results can be calculated directly by the computer program, which is not possible in traditional method in which a subject must write down the elements he or she heard out. As can be seen this method allows to avoid also typo and spelling mistakes. The panel contains the words that were present in the sentence and some paronyms (words that sound similarly and can be confused with the correct ones). The detailed description of the method can be found in (OZIMEK *et al.*, 2009c). After each correct answer the level of the next sentence was decreased, while after each incorrect answer the level of the next sentence was increased. The level of the babble noise was constant and equal to 70 dB SPL. In this way each sentence was presented at different SNR but the procedure led to the threshold of 50%

correct answers. One list was used to determine each SRT, which was calculated as a mean value from the last eight (from thirteen in total) SNRs (OZIMEK *et al.*, 2009c).

To get the reference value of the SRT, the test was also recorded via a dummy head in an anechoic chamber. The signal source was placed at the same position as noise source (i.e. in front of the dummy head). Then the listening session for 20 normal hearing subjects using the standard adaptive procedure was carried out. The average SRT for this condition was determined ( $SRT = -5.6$ ) and taken as a reference value. This method is very fast (lasts 3–4 minutes) and automated.

For both tests, the so-called binary scoring was used in the assessment of speech recognition (OZIMEK *et al.*, 2009c). Only a correctly written logatome/sentence was counted as correctly understood and any mistake (except spelling mistakes) led to

an incorrect note. All subjects were paid and took part in a short training session to familiarize with the task. For each measurement point in a given enclosure, results from ten normal hearing subjects were collected; thus, each SRT was the average of ten individual values.

#### 4. Objective parameters in the tested enclosures

In each room the measurements of the objective parameters were carried out in many measurement points. The ranges of the values of the objective parameters (RT, EDT, weighted  $C_{50}$ , STI) obtained in the six tested enclosures (church, assembly hall of a music school, two courtrooms of different volumes, a typical auditorium and a university hall) are given in Table 1. Since all the parameters were measured in many points it is a big number of individual data, thus they

Table 1. Range of objective parameters and speech recognition parameters values measured in tested enclosures. The RT and EDT and  $C_{50}$  were measured according to ISO 3382 (2010), then  $C_{50}$  was weighted according to (MARSHALL, 1994). STI was determined according to IEC 60268-16 norm (2011).

Room	Parameter	Range
Church	RT [s]	$\langle 3.6, 4.2 \rangle$
	EDT [s]	$\langle 3.1, 4.6 \rangle$
	weighted $C_{50}$	$\langle -39.9, -1.4 \rangle$
	STI	$\langle 0.28, 0.58 \rangle$
	Background noise [dB(A)]	37
Adam Mickiewicz University Hall	RT [s]	$\langle 1.59, 1.98 \rangle$
	EDT [s]	$\langle 1.77, 1.87 \rangle$
	weighted $C_{50}$	$\langle -5, 5, -1.8 \rangle$
	STI	$\langle 0.5, 0.63 \rangle$
	Background noise [dB(A)]	40
Large courtroom	RT [s]	$\langle 0.88, 0.98 \rangle$
	EDT [s]	$\langle 0.70, 1.07 \rangle$
	weighted $C_{50}$	$\langle -2.1, 8.6 \rangle$
	STI	$\langle 0.58, 0.8 \rangle$
	Background noise [dB(A)]	34
Auditorium	RT [s]	$\langle 0.60, 0.66 \rangle$
	EDT [s]	$\langle 0.46, 0.65 \rangle$
	weighted $C_{50}$	$\langle -1.9, 7.3 \rangle$
	STI	$\langle 0.47, 0.8 \rangle$
	Background noise [dB(A)]	38
Assembly hall	RT [s]	$\langle 1.06, 1.20 \rangle$
	EDT [s]	$\langle 0.93, 1.11 \rangle$
	weighted $C_{50}$	$\langle -1.9, 2.2 \rangle$
	STI	$\langle 0.57, 0.64 \rangle$
	Background noise [dB(A)]	34
Small courtroom	RT [s]	$\langle 0.78, 0.81 \rangle$
	EDT [s]	$\langle 0.68, 0.75 \rangle$
	weighted $C_{50}$	$\langle 1.1, 5.1 \rangle$
	STI	$\langle 0.65, 0.72 \rangle$
	Background noise [dB(A)]	34

Table 2. Averaged across subjects range of speech recognition parameters in particular enclosures.

Room	Speech recognition parameter	Range
Church	Logatome recognition [%]	⟨24, 65⟩
	SRT of sentence test in noise [dB]	⟨−7.1, 6.6⟩
Adam Mickiewicz University Hall	Logatome recognition [%]	⟨59, 72⟩
	SRT of sentence test in noise [dB]	⟨−3.1, −4.9⟩
Large courtroom	Logatome recognition [%]	⟨65, 86⟩
	SRT of sentence test in noise [dB]	⟨−8.9, −3.5⟩
Auditorium	Logatome recognition [%]	⟨43, 82⟩
	SRT of sentence test in noise [dB]	⟨−0.5, −7.3⟩
Assembly hall	Logatome recognition [%]	⟨62, 71⟩
	SRT of sentence test in noise [dB]	⟨−3.5, −6.7⟩
Small courtroom	Logatome recognition [%]	⟨79, 86⟩
	SRT of sentence test in noise [dB]	⟨−6.9, −8.3⟩

are not presented here. Moreover, the main purpose of this study is to find the relationship between those parameters and speech recognition. It must be emphasized that individual data for each measurement point in each enclosure, so values of objective parameter as well as values of speech recognition were taken to the analysis carried out in Sec. 5. Moreover, while there are a lot of rooms with relatively short RT (below 1 s, e.g. auditoria) and relatively long RT (over 3 s, e.g. churches), the enclosures with RT around 2 s are rarer (mainly classical concert halls). Thus, there is a small gap in the measures values of RT between 2 s and 3.5 s.

### 5. Logatome and sentence recognition in the tested enclosures

The values of the logatome and sentence recognition obtained in the six tested enclosures are given in Table 2. Again, not to obscure the main purpose of the study, only mean values (across subjects) are given here, while each measurement point separately was taken into account for further analysis (see Sec. 6).

### 6. Relationship between objective measures and speech recognition

The most important aspect considered in this paper is the relationship between objective measures of the tested enclosures and speech recognition. Figure 1 shows the logatome recognition as a function of reverberation time for all investigated enclosures and all measurement points in each of the rooms. Speech recognition for each measurement point was averaged across subjects. As can be seen there is a general dependency showing that for short RTs, the recognition is high, while it decreases with the increase of RT. The same tendency can be noticed for EDT (Fig. 2).

Linear fitting was used to describe the relationship between those values with  $R^2 = 0.75$  (for RT)

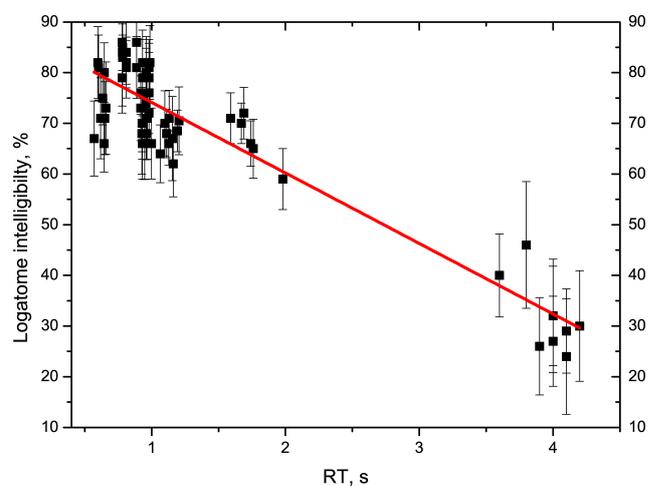


Fig. 1. Logatome recognition vs. RT for different enclosures.

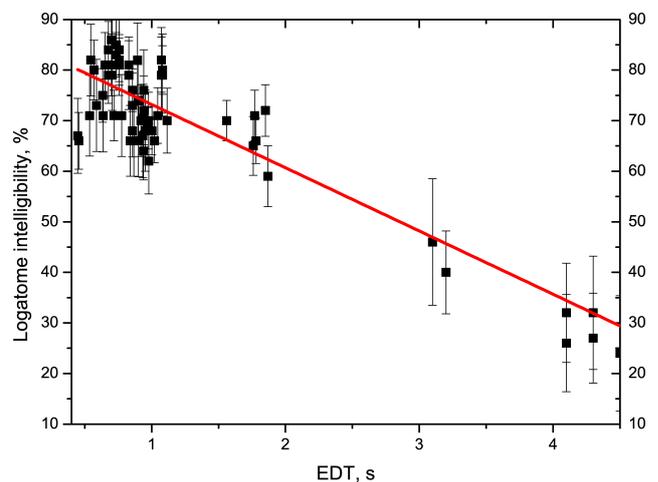


Fig. 2. Logatome recognition vs. EDT for investigated enclosures.

and  $R^2 = 0.71$  (for EDT). The relationships between Logatome intelligibility (LI) and RT or EDT are as follows:

$$LI = -13.9RT + 88, \quad LI = -12.5EDT + 85.7.$$

As can be seen there is a significant correlation between logatome recognition and RT and EDT.

Figure 3 shows logatome recognition as a function of weighted  $C_{50}$ .

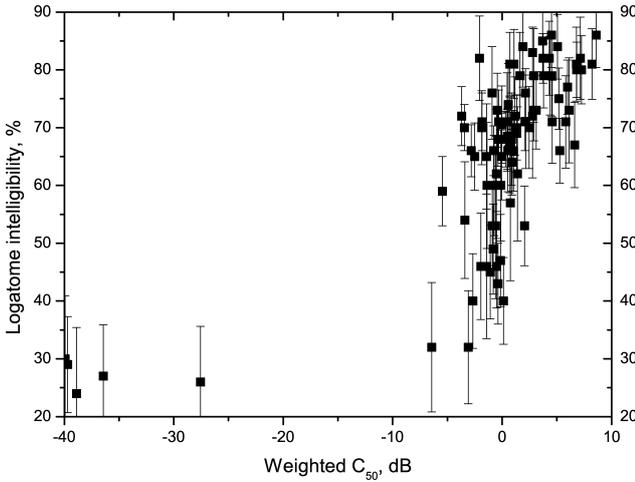


Fig. 3. Logatome recognition vs. weighted  $C_{50}$  for all investigated enclosures.

No function could be fitted to the obtained data with the  $R^2$  greater than 0.52, thus no fitting is shown, nevertheless, a general conclusion is that for low values of  $C_{50}$  parameter, the recognition is low and it increases along with the increase of  $C_{50}$ .

The most important relationship, however, is logatome recognition versus STI, which is shown in Fig. 4.

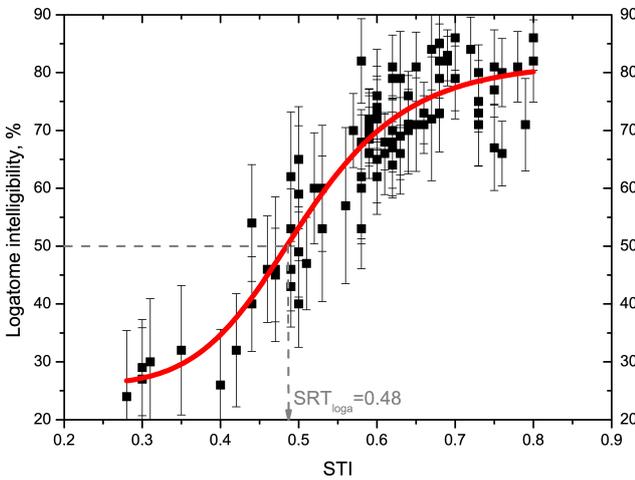


Fig. 4. Logatome recognition vs. STI for investigated enclosures. The SRT for logatomes in terms of STI is also given here ( $SRT_{log a}$ ).

A logistic function according to (GARCÍA-PÉREZ *et al.*, 2011; LAM *et al.*, 1997; OZIMEK *et al.*, 2009b; 2009c; SHEN, RICHARDS, 2012) was fitted to the data ( $R^2 = 0.78$ ). Parameters for fitted curve are as follows:

$$LI = \frac{(25.98 - 82)^{7.33}}{1 + \left(\frac{STI}{0.5}\right)} + 82.$$

Again, as can be seen there is a strong correlation between logatome recognition and STI.

Similar relations can be drawn for speech reception threshold (SRT expressed in S/N ratio (in dB)) and objective parameters for tested enclosures. Figures 5 and 6 show the relationship between SRT and RT and EDT, respectively. It can be generally noticed that the SRTs in noise are relatively low for short RTs (EDTs) (recognition is high), while SRTs increase with the increase of RT (EDT) which is in line with DUQUESNOAYN and PLOMP research (1980).

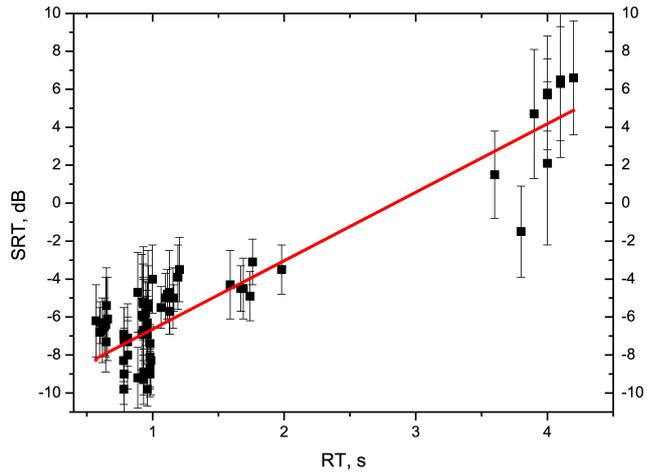


Fig. 5. SRT vs. RT for investigated enclosures.

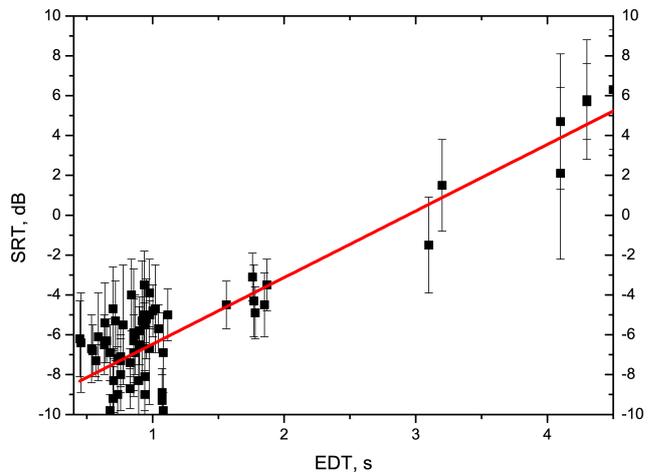


Fig. 6. SRT vs. EDT for investigated enclosures.

Similarly to the logatomes, linear fitting was used to derive the relationship between SRT and RT ( $R^2 = 0.7$ ) or EDT ( $R^2 = 0.7$ ). Those functions are given below:

$$SRT = 3.6RT - 10.3, \quad SRT = 3.3EDT - 9.8.$$

Figure 7 shows the relationship between SRT and  $C_{50}$  values. A general finding is that the value of  $C_{50}$  from the range  $-5$  to  $10$  dB, influences the SRT, however no curve can be fitted to the results.

The relationship between SRT and STI is depicted in Fig. 8.

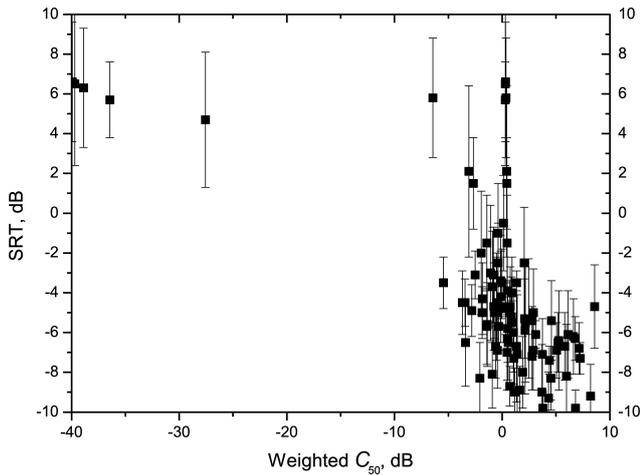
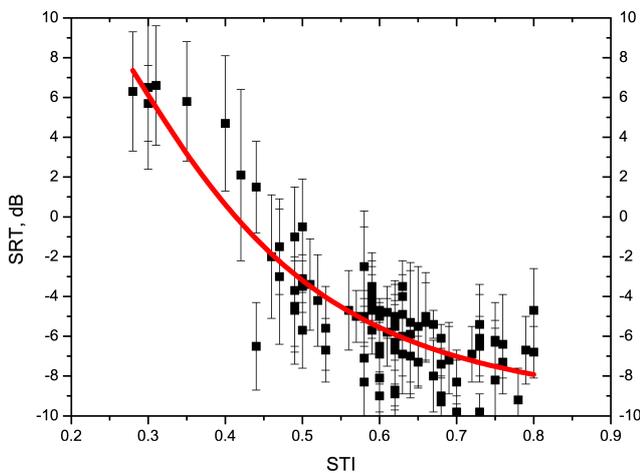
Fig. 7. SRT vs.  $C_{50}$  for investigated enclosures.

Fig. 8. SRT vs. STI for investigated enclosures.

Similarly to logatomes, a logistic curve was fitted to the data with  $R^2 = 0.6$ :

$$\text{SRT} = \frac{(16.28 + 10.05)}{1 + \left(\frac{\text{STI}}{0.35}\right)^{2.95}} - 10.05.$$

The correlation coefficient in this case is somewhat smaller than that obtained for logatome recognition. As can be seen from Fig. 8, sentence speech recognition determined by SRT is nonlinear function of STI. The most sensitive range of SRT relative to STI (the steepest slope of the psychometric function (slope about 10 dB SRT per 0.3 STI)) covers the range of STI changes from 0.3 to 0.6. Below and above this range, SRT is much less sensitive to STI changes. As there is not much data for  $\text{STI} < 0.3$ , one may expect that for such bad conditions SRT value would asymptotically tend to infinity, which is in line with (HOUTGAST, STEENEKEN, 2002). On the other hand, for  $\text{STI} > 0.6$  the SRT values are the same as for anechoic conditions which can be described as the best ones. This statement is the new and main finding resulting from the

present study. Similar conclusion can be drawn from Fig. 4 referring to logatome recognition versus STI.

Finally, the relationship between SRT values and logatome recognition can be found using the same procedure of logistic function fitting ( $R^2 = 0.7$ ). The results are given in Fig. 9

$$\text{SRT} = \frac{(71.25 - 157.12)}{1 + \left(\frac{\text{LI}}{1294.99}\right)^{0.23}} + 157.12.$$

As can be noticed a relatively high correlation was found.

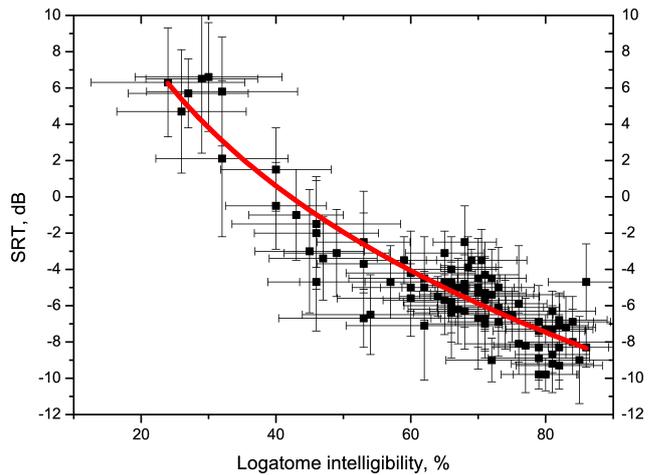


Fig. 9. SRT vs. logatome recognition for investigated enclosures.

## 7. Discussion

In this study both objective parameters and subjective measures of speech recognition were derived. Comparing logatome recognition and sentence recognition expressed by SRT data it can be noticed that the results are generally consistent. On the basis of the data obtained one may say that the proposed procedure of SRT determination can be used instead of classical speech recognition measurements. This procedure is faster and fully automated. What is more it is based on the sentences used in everyday communication. Thus, it is reasonable to state that it reflects in a most direct and robust way the real communication process which is mainly based on sentences. It also gives smaller inter-individual differences.

Comparing the obtained results with the data gathered by other authors it must be stated that the weighted  $C_{50}$  proposed by MARSHALL (1994) does not reflect speech perception process in a proper way. Comparing logatome intelligibility and sentence SRT with  $C_{50}$  it may be stated that  $C_{50}$  gives equivocal results for different reverberation conditions for both relationships. Thus, it cannot be used as a predictor of speech intelligibility. This fact was found in both logatome recognition and SRT determination. Much stronger relation was found with the simple parameters such as

RT and EDT which seem to be linear in the analyzed range of values.

Other fitting was based on the logistic function as this dependency is the most common in psychoacoustic parameters in which psychometric function is determined. It must be emphasized that this shape of curve gives the highest  $R^2$  values from the large range of different fitting curves tested.

The obtained logatome recognition vs. STI is generally consistent with the data presented by HOUTGAST and STEENEKEN (2002) for  $STI < 0.4$ . Nevertheless, some shifts can be observed for higher STI values. The logatome SRT ( $SRT_{\log a}$ ) value in terms of STI, calculated accordingly to (HOUTGAST, STEENEKEN, 2002) in this study is obtained for  $STI = 0.48$  (see Fig. 4 for details), while HOUTGAST and STEENEKEN suggest that it should be around  $STI = 0.38$ . Moreover, the slope for logatome vs. SRT is about 17.5%/0.1 units (it means that with the STI increase of 0.1 units the logatome recognition increases of about 17.5%), whereas for the curve presented by HOUTGAST and STEENEKEN the slope is about 22.5%/0.1 units which is 5%/0.1 higher. Therefore, for higher STI values there is a higher shift of the presented results comparing to the Houtgast and STEENEKEN than for lower STIs. This findings are in line with the findings of BRACHMAŃSKI (2004; 2015).

The preliminary research in this area has confirmed that a use of sentence test brings equivocal results since 100% recognition is obtained even for very bad conditions ( $STI > 0.4$ ). This is caused by the high context effect. This problem was eliminated by using masker. It must be emphasized, however, that the results of a proposed method reflect the convolutive distortions (reverberation) only which are the most important in the enclosure, and the slight effect of distance (which should be taken into account in very large enclosures) is neglected which can be somewhat reflected in the results by a light lower SRTs (higher recognition). This effect will be analyzed in the future work.

## 8. Conclusions

On the basis of the results obtained in the study the following main conclusions can be drawn:

- The study provides new data in the subjective assessment of speech intelligibility in rooms, especially in terms of SRT determination using sentence tests.
- Comparison of the data gathered for sentence tests (SRT) and logatome tests (LI), gave the insight into the relation between those two values. A mathematical relationship between them was derived.
- The method introduced here, which is automated and fast, can be used in many cases instead of constant stimuli one.

- The validation of STI was carried out using new method based on SRT determination using sentence tests.
- Weighted  $C_{50}$  according to MARSHALL (1994) was proven to be poor speech intelligibility predictor.

## Acknowledgments

Supported by a grant from Norway through the Norwegian Financial Mechanism (project no. PNRF-167-AI-1/07) and a grant from National Science Centre: Project Number UMO-2011/03/B/HS6/03709. The authors thank Michał Sołoducha for technical assistance.

## References

1. BORISH J., ANGELL J. (1983), *An Efficient Algorithm for Measuring the Impulse Response Using Pseudorandom Noise*, Journal of the Audio Engineering Society, **31**, 7/8, 478–488.
2. BRACHMAŃSKI S. (2004), *Estimation of logatome intelligibility with the STI method for Polish speech transmitted via communication channels*, Archives of Acoustics, **29**, 4, 555–562.
3. BRACHMAŃSKI S. (2008), *Objective measure for assessment of speech quality in rooms*, Archives of Acoustics, **33**, 4(S), 177–182.
4. BRACHMAŃSKI S. (2015), *Chosen problems of speech transmission quality assessment* [in Polish: *Wybrane zagadnienia oceny jakości transmisji sygnału mowy*], Oficyna Wydawnicza PWr, Wrocław.
5. BRACHMAŃSKI S., STARONIEWICZ P. (1999), *Phonetic structure of a test material used in subjective measurements of speech quality* [in Polish: *Fonetyczna struktura materiału testowego stosowanego w subiektywnych pomiarach jakości mowy*], Speech and Language Technology, **3**, pp. 71–80.
6. BRADLEY J.S. (1986), *Predictors of speech intelligibility in rooms*, Journal of Acoustical Society of America, **80**, 837–845.
7. BRADLEY J.S., SATO H., PICARD M. (2003), *On the importance of early reflections for speech in rooms*, Journal of Acoustical Society of America, **113**, 3233–3244.
8. CHU W.T. (1990), *Impulse-response and reverberation-decay measurements made by using a periodic pseudo-random sequence*, Applied Acoustics, **29**, 3, 193–205.
9. DUQUESNOY A.J., PLOMP R. (1980), *Effect of reverberation and noise on the intelligibility of sentences in cases of Presbycusis*, Journal of the Acoustical Society of America, **68**, 2, 537–544.
10. GARCÍA-PÉREZ M.A., ALCALÁ-QUINTANA R., WOODS R.L., PELI E. (2011), *Psychometric functions for detection*

- and discrimination with and without flankers, *Attention Perception Psychophys*, **73**, 829–853.
11. GEORGE E.L., GOVERTS S.T., FESTEN J.M., HOUTGAST T. (2010), *Measuring the effects of reverberation and noise on sentence intelligibility for hearing-impaired listeners*, *Journal of Speech Language and Hearing Research*, **53**, 6, 1429–1439.
  12. HAGERMAN B. (1982), *Sentences for testing speech intelligibility in noise*, *Scandinavian Audiology*, **11**, 79–87.
  13. HOUSE A.S., WILLIAMS C.E., HECKER M.H.L., KRYTER K.D. (1965), *Articulation testing methods: Consonantal differentiation with a Closed-Response set*, *Journal of Acoustical Society of America*, **37**, 1, 158–166.
  14. HOUTGAST T., STEENEKEN H.J. (1985), *A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria*, *Journal of Acoustical Society of America*, **77**, 1069–1077.
  15. HOUTGAST T., STEENEKEN H.J. (2002), *Past, present, and future of the Speech Transmission Index*, TNO Human Factors. Soesterberg, The Netherlands.
  16. HOUTGAST T., STEENEKEN H.J.M. (1973), *The modulation transfer function in room acoustics as a predictor of speech intelligibility*, *Acta Acustica United with Acustica*, **28**, 66–73.
  17. HOUTGAST T., STEENEKEN H.J.M., PLOMP R. (1980), *Predicting speech intelligibility in rooms from the modulation transfer function. I. General room acoustics*, *Acustica*, **46**, 60–72.
  18. IEC60268-16 (2011), *Sound system equipment. Part 16: Objective rating of speech intelligibility by speech transmission index*.
  19. JACOB K.D., BIRKLE T.K., ICKER C.B. (1991), *Accurate Prediction of speech intelligibility without the use of in-room measurements*, *Journal of the Audio Engineering Society*, **39**, 4, 232–242.
  20. KALIKOW D.N., STEVENS K.N., ELLIOT L.L. (1977), *Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability*, *Journal of Acoustical Society of America*, **61**, 1337–1351.
  21. KOLLMEIER B., WESSELKAMP M. (1997), *Development and evaluation of a German Sentence Test for objective and subjective Speech Intelligibility Assessment*, *Journal of Acoustical Society of America*, **102**, 4, 2412–2421.
  22. KUTTRUFF M. (2009), *Room Acoustics*, 5th Ed., Spon Press. Abington, Oxon, UK.
  23. LAM C.F., DUBNO J.R., AHLSTROM J.B., HE N.J., MILLS J.H. (1997), *Estimating parameters for psychometric functions using the four-point sampling method*, *Journal of the Acoustical Society of America*, **102**, 6, 3697–3703.
  24. LEVITT H. (1971), *Transformed up-down methods in psychoacoustics*, *Journal of Acoustical Society of America*, **49**, 467–477.
  25. LONGWORTH-REED L., BRANDEWIE E., ZAHORIK P. (2008), *Time-forward speech intelligibility in time-reversed rooms*, *JASA Express Letters*, doi: 10.1121/1.3040024,
  26. MAJEWSKI W., MYŚLECKI W., BAŚCIUK K., BRACHMAŃSKI S. (1998), *Application of modified logatome intelligibility test in telecommunications, audiometry and room acoustics*, 9th Mediterranean Electrotechnical Conference Melecon'98, Tel-Aviv, Israel, 25–28.
  27. Bruel & Kjaer (2008), *Manual Dirac 4.1*.
  28. MARSHALL L.G. (1994), *An acoustic measurement program for evaluating auditoriums based on the early/late sound energy ratio*, *Journal of Acoustical Society of America*, **96**, 4, 2251–2261.
  29. NILSSON M., SOLI S.D., SULLIVAN J.A. (1994), *Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise*, *Journal of Acoustical Society of America*, **95**, 1085–1099.
  30. OZIMEK E., KUTZNER D., LIBISZEWSKI P., WARZYBOK A., KOCINSKI J. (2009a), *The new polish tests for speech intelligibility measurements*, 13th IEEE SPA, Signal Processing: Algorithms, Architectures, Arrangements, and Applications, Poznań.
  31. OZIMEK E., KUTZNER D., SEK A., WICHER A. (2009b), *Development and evaluation of Polish digit triplet test for auditory screening*, *Speech Communication*, **51**, 4, 307–316.
  32. OZIMEK E., KUTZNER D., SEK A., WICHER A. (2009c), *Polish sentence tests for measuring the intelligibility of speech in interfering noise*, *International Journal of Audiology*, **48**, 7, 433–443.
  33. OZIMEK E., KUTZNER D., SEK A.P., WICHER A., SZCZEPANIAK O. (2006), *The Polish sentence test for speech intelligibility evaluations measurements*, *Archives of Acoustics*, **31**, 4, 431–438.
  34. PENG J. (2008), *Relationship between Chinese speech intelligibility and speech transmission index in rooms using dichotic listening*, *Chinese Science Bulletin*, **53**, 18, 2748–2752.
  35. PLOMP R., MIMPEN A.M. (1979), *Improving the reliability of testing the speech reception threshold for sentences*, *Audiology*, **18**, 43–53.
  36. PN-EN-ISO3382 (2010), *Acoustics – Measurement of room acoustic parameters – Part 2: Reverberation time in ordinary rooms*.
  37. PRUSZEWICZ A., DEMENKO G., RICHTER L., WIKI T. (1994a), *New articulation lists for speech audiometry. Part I [in Polish]*, *Otolaryngologia Polska*, **48**, 50–55.

38. PRUSZEWICZ A., DEMENKO G., RICHTER L., WIKI T. (1994b), *New articulation lists for speech audiometry. Part II* [in Polish], *Otolaryngologia Polska*, **48**, 56–62.
39. SHEN Y., RICHARDS V.M. (2012), *A maximum-likelihood procedure for estimating psychometric functions: Thresholds, slopes, and lapses of attention*, *Journal of the Acoustical Society of America*, **132**, 2, 957–967.
40. STEENEKEN H.J.M., HOUTGAST T. (1980), *A physical method for measuring speech-transmission quality*, *Journal of the Acoustical Society of America*, **69**, 318–326.
41. VERSFELD N.J., DAALDER L., FESTEN J.M., HOUTGAST T. (2000), *Method for the selection of sentence material for efficient measurement of the speech reception threshold*, *Journal of Acoustical Society of America*, **107**, 1671–1684.
42. YANG W. (2006), *Optimizing Acoustical conditions for speech intelligibility in classrooms*, PhD Thesis, University of British Columbia, Vancouver.
43. YANG W., BRADLEY J.S. (2009), *Effects of room acoustics on the intelligibility of speech in classrooms for young children*, *Journal of Acoustical Society of America*, **125**, 2, 922–933.