AUTOMATION OF THE LOGATOM INTELLIGIBILITY MEASUREMENTS IN ROOMS

Stefan BRACHMAŃSKI

Wrocław University of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
e-mail: Stefan.Brachmanski@pwr.wroc.pl

Speech intelligibility is one of basic quality parameters of speech transmission in rooms. The methods for assessment of speech quality fall into two classes: subjective and objective methods. This paper includes an overview of selected methods of subjective listening measurements (ACR – Absolute Category Rating, DCR – Degradation Category Rating, speech intelligibility) recommended by ITU-T, ISO and Polish Standard and the method of speech transmission quality evaluation called "modified intelligibility test with forced choice" (MIT-FC). The MIT-FC method provides fully automatized measurement of speech intelligibility in rooms. The experiments carried out in finding the relations between logatom intelligibility measured with traditional and the MIT-FC methods for the rooms have shown that there exists the multivalue and repetitive relation between them.

Keywords: speech quality, speech intelligibility, room acoustic.

1. Introduction

One of important elements of communications is a quality of transmission which depends on objective, or physical, parameters of rooms as well as on the subjective factors connected to listeners in that room. Measurements of speech transmission quality should take into account some subjective factors by the use of subjective measurements methods, or by the estimation of subjectively weighted objective results. Among the different subjective methods, the techniques which give directly [7, 4] or indirectly [1–3, 5, 6, 8] the values of Mean Opinion Score (MOS) on the five-grade quality scale are used.

2. Absolute category rating

The ACR (*Absolute Category Rating*) method is recommended by ITU [7] for the evaluation of subjective quality of the speech. The speech material (test lists) used in this method should consist of simple, short, semantically unrelated sentences. A test list is divided into groups of five sentences. The test material should be properly prepared and

recorded. The speaker should pronounce the sentences fluently and should not have any speech defects. Since the female voice and the male voice have different characteristics, the two types of voice should be included in the measurements. The results obtained for male and female voices should be evaluated separately. They can be averaged only when they do not differ significantly. To reduce the influence of the individual characteristics of the speaker's voice on the obtained result, several speakers should take part in the experiment. The experiment's listening part should take place in a room with a noise level below 30 dBA. Listeners are chosen at random from the normal telephone using population, with the provisions that:

- they have not been involved in work connected with assessment of performance of telephone systems or speech coding,
- they have not participated in any subjective measurements at least the previous six months,
- they have never heard the same sentences lists before.

Listeners listen to the sentences and give their opinions in five levels scale. Various scales recommended by ITU may be used for different purposes:

- listening-quality scale (Excellent speech is rated 5, Good – 4, Fair – 3, Poor – 2, Bad – 1),
- listening-effort scale (Complete relaxation possible; no effort required is rated 5, Attention necessary; no appreciable effort required – 4, Moderate effort required – 3, Considerable effort required – 2, No meaning understood with any feasible effort – 1),
- loudness-preference scale (Much louder than preferred is rated 5, Louder than preferred – 4, Preferred – 3, Quieter than preferred – 2, Much quieter than preferred – 1).

The average rating (Mean Opinion Score – MOS) is calculated over the listeners and the speakers for each tested speech transmission condition.

3. The traditional method of logatom intelligibility measurement

Subjective tests are described in Polish Standard PN 90/T-05100 "Analog Telephone Chains. Requirements and Methods of Measuring Logatom Articulation".

The measurement of logatom[1] intelligibility consists in the transmission of logatom lists, read out by a speaker, through the tested channel, which are then written down by listeners and the correctness of the record is checked by a group of experts who calculate the average logatom intelligibility. It is recommended to use lists of 50 or 100 logatoms. Each list should be phonetically and structurally balanced. The measurement should be carried out in rooms in which level of internal noise together with external noise (not introduced on purpose) does not exceed 40 dBA. The listeners should be selected from persons who have normal, good hearing and normal experience in pronunciation

---

[1] Logatom – (*logos* (gr.) – spoken phrase, *atom* (gr.) – indivisible) vocal sound, generally insignificant, usually made by the sound of a consonant or the first consonant, then by an intermediate vowel, finally by a consonant or a final consonant sound.

in the language used in the test. A person is considered to have normal hearing if her/his threshold does not exceed 10 dB for any frequency in a band of 125 Hz–4000 Hz and 15 dB in a band of 4000 Hz–6000 Hz. Hearing threshold should be tested by means of a diagnostic audiometer. The size of the listening group should be such that the obtained averaged test results do not change as the group size is further increased (minimum 5 persons). The group of listeners who are to take part in logatom intelligibility measurements should be trained (2–3 training sessions are recommended). Logatoms should be spoken clearly and equally loudly without accenting their beginnings or ends. The time interval between individual logatoms should allow the listener to record the received logatom at leisure. It is recommended that logatoms should be spoken with 3–5 sec pauses in between. The time interval between sessions should not be shorter than 24 h and not longer than 3 days. The total duration of a session should not exceed 3 hours (including 10 minute breaks after each 20 minute listening period).

Listeners write the received logatoms on a special form on which also the date of the test, the test list number, the speaker's name or symbol (no.), the listener's name and additional information which the measurement manager may need from the listener is noted. The recording should be legible to prevent a wrong interpretation of the logatom. The received logatoms may be written in phonetic transcription (a group of specially trained listeners is needed for this) or in an orthographic form specific for a given language. In the next step, the group of experts checks the correctness of received logatoms and the average logatom intelligibility is calculated in accordance to the Eqs. (1) and (2):

$$W_L = \frac{1}{N \cdot K} \sum_{n=1}^{N} \sum_{k=1}^{K} W_{n,k} \ \ [\%], \tag{1}$$

N – number of listeners, K – number of test lists, $W_{n,k}$ – logatom intelligibility for n-th listener and k-th logatom list,

$$W_{n,k} = \frac{P_{n,k}}{T_k} \cdot 100 \ \ [\%], \tag{2}$$

$P_{n,k}$ – number of correctly received logatoms from k-th logatom list by n-th listener, T_k – number of logatoms in k-th logatom list.

4. Modified intelligibility test with forced choice (MIT-FC)

The subjective measurement of logatom intelligibility is very time- and cost-consuming. To avoid disadvantages of subjective evaluation of logatom intelligibility by means of the "traditional" method, a new measurement method was created and developed at the Institute of Telecommunications, Teleinformatics and Acoustics. This method was called "*modified intelligibility test with forced choice*" *(MIT-FC)*.

In the MIT-FC method all experiments are controlled by a computer. The automation of the subjective measurement is connected with the basic change in generation of logatoms and in making decision by a listener. The computer generates logatoms

and presents the utterances, via a D/A converter and loudspeaker to the listeners subsequently and for each spoken utterance several logatoms that have been previously selected as perceptually similar are visually presented. It has been found that the optimal number of logatoms presented visually to the listeners is seven (six alternative logatoms and one transmitted logatom to be recognized). The listener chooses one logatom from the list visually presented on the computer monitor. The computer counts the correct answers and calculates the average logatom intelligibility and standard deviation.

5. Experiments

The goals of experiment:
- decision if the results of traditional and modified with forced choice methods let finding the relation which would allow to convert results from one method to the other and the classification of rooms tested with both methods,
- measurement of experimental relations between traditional and modified logatom intelligibility methods.

The subjective tests were done according to Polish Standard PN-90/T-05100 [8] and Recommendation ITU-T P.800 [7] with the team of listeners made up of 12 listeners in age from 18 to 25 years. The listening team was selected from persons with normal hearing. The qualification was based on audiometric tests of hearing threshold. The measurements of logatom intelligibility were done using the traditional method and the MIT-FC method.

The measurements were taken in two unoccupied rooms. In each room, four listener locations were selected. These positions were chosen in the expectation of yielding a wide range of logatom intelligibility. Sound sources (voice and white noise) were positioned in the part of the room normally used for speaking. One loudspeaker was the voice source and the second – the noise source. The various conditions were obtained by combination five level of white noise.

The testing material consisted of phonetically and structurally balanced logatoms and sentences lists uttered by professional male speaker, whose native language was Polish. For each measure point (the place where the measure position was situated) a list of 100 logatoms has been prepared. The logatom lists at the four listener locations were recorded on the digital tape recorder. These recordings were played back over headphones to the subject afterward.

This way of subjective measurements realization provides the same listening conditions for both traditional and with choice methods. In each room for each position of listener (Pp) and for each signal-to-noise ratio (SNR) the logatom intelligibility was obtained by averaging out the group of listeners results.

The results of subjective measurements of logatom intelligibility are shown in Fig. 1. After the logatom intelligibility measurements, the listeners assessed the quality of speech transmission in range from 1 to 5 according to the MOS speech quality scale. The obtained results are partially presented in Table 1. In this table the values of MOS
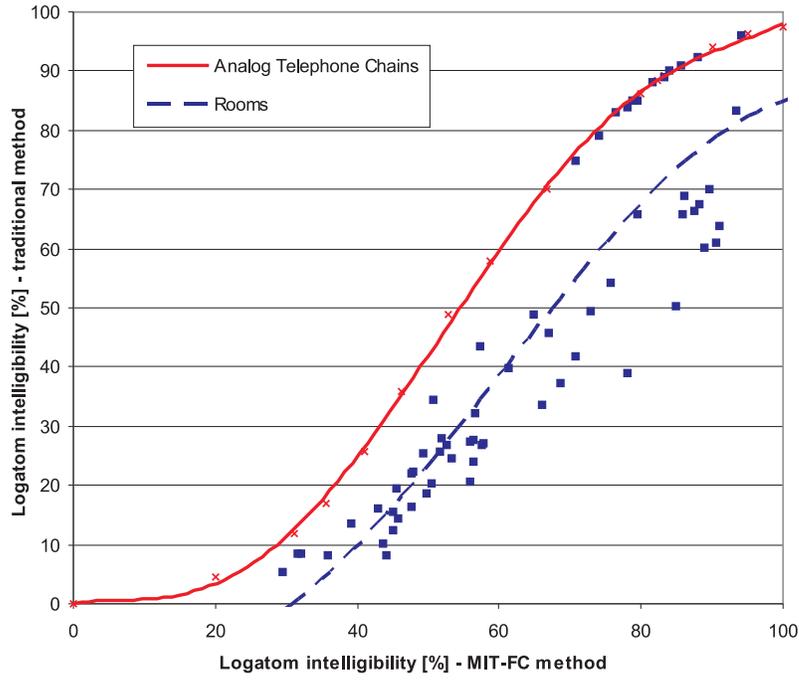
Fig. 1. Relationship between logatom intelligibility measured with traditional and MIT-FC method for analog telephone chains and rooms.

Table 1. Logatom intelligibillity and MOS (ACR) of auditoria measurements.

SNR	MIT-FC				Traditional method					MOS_Wl	MOS_ACR
	Pp1	Pp2	Pp3	Pp4	Pp1	Pp2	Pp3	Pp4	Wl		
0	41	45.5	45.5	49.8	14.7	15.6	19.5	18.7	17.13	1	1
3	49.4	56	51.8	49.2	23.2	20.7	28.1	25.4	24.35	1	1
6	53.2	57.9	62	50.6	25.8	25.1	43.5	34.4	32.20	1.3	1.4
9	56.6	70.8	64.8	61.4	32.2	34.4	49	39.8	38.85	1.6	2
12	65.3	75.2	80.9	75.6	45.8	47.4	65.9	55.2	53.58	2.5	3
15	84.33	85.2	85.4	85	36.33	56.66	62.33	45.5	50.21	2.2	3
18	88.2	88.25	91.2	88.2	68	56.66	65	51	60.17	3	3.2
21	83	86	87.2	88	64.33	63.5	73.5	74	68.83	3.6	3.4
24	90	92.2	90.2	90.2	69.33	56.25	61.5	56.5	60.90	3	3.5
27	85.33	84.33	87.4	86.6	64.75	56.8	68.25	74	65.95	3.4	3.6
30	88.5	84.5	91.2	89	64.25	62.25	76	67.25	67.44	3.5	3.9
33	88.25	85.66	88.25	87.5	63.25	61	68.75	72.33	66.33	3.4	3.9
36	89.66	87.33	89	92.33	66.4	60	79	74.6	70.00	3.4	4
39	90.33	86	93.8	94	59.67	60	70.5	65.5	63.92	3.3	4.1

and quality standards, obtained on the basis of the data given in Polish Standard (PN-90/T-05100), are also presented.

6. Conclusion

The experiments carried out in finding the relations between logatom intelligibility measured with traditional and semi-automatic with forced choice methods for the rooms have shown that there exist the relation between them. It allows using both methods interchangeably and converting results between them.

The presented MIT-FC method offers a simple, easy to use, stable, and fully automatized speech system to assessment of speech quality in rooms. The results of the experiments have shown that the MIT-FC method is very useful in the evaluation of speech quality in rooms. The time needed to carry out the measurement with MIT-FC method is the same as in traditional one but we obtain the results right after finishing the measurement process.

The results of the presented experiments are the first step in the subjective assessment of speech quality in rooms research. The next stage is the realization of subjective measurements with both methods with considering other kinds of distortion which can occur in rooms.

References

[1] BASCIUK K., BRACHMAŃSKI S., *The automation of the subjective measurements of logatom intelligibility*, 102-nd Convention AE S, Munich, Prep. 4407, 1997.

[2] BRACHMAŃSKI S., *Assessment of Quality of Speech Transmitted over IP Networks*, Internet Technologies, Applications and Societal Impact, WITASI 2002, pp. 1–14, Kluwer Academic Publishers, 2002.

[3] BRACHMAŃSKI S., *The automation of subjective measurements of speech intelligibillity in rooms*, The 112th Conv. AES, Monachium, Preprint 5588, 2002.

[4] BRACHMAŃSKI S., *Experimental comparison between speech transmission index (STI) and mean opinion scores (MOS) in rooms*, Arch. Acoust., **31**, 4, 171–176 (2006).

[5] DAVIES D. D., DAVIES C., *Application of speech intelligibility to sound reinforcement*, J. Audio Eng. Soc., **37**, 12, 1002–1018 (1989).

[6] MACKIE K., *Assessment of evaluation measures for processed speech*, Speech Comm., **6**, 309–316 (1987).

[7] ITU-T Rec P.800, *Method for subjective determination of transmission quality*, 1996.

[8] PN-T-05100, *Analogowe łańcuchy telefoniczne. Wymagania i metody pomiaru wyrazistości logatomowej*, Polska Norma.