ARCHIVES OF ACOUSTICS
29, 4, 555–562 (2004)

ESTIMATION OF LOGATOM INTELLIGIBILITY WITH THE STI METHOD FOR POLISH SPEECH TRANSMITTED VIA COMMUNICATION CHANNELS

S. BRACHMAŃSKI

Wrocław University of Technology
Institute of Telecommunications and Acoustics
Wybrzeże Wyspiańskiego 24, 50-370 Wrocław, Poland
email: stefan.brachmanski@pwr.wroc.pl

(*received 23 February 2004; accepted 23 September 2004*)

An objective measure of the quality of speech communication channels was introduced and described by Steeneken- Houtgast. By calculating the Speech Transmission Index (STI), this method provides an estimate of logatom intelligibility for a variety of communication distortions. Using the logatom test for speech intelligibility, it was shown that the STI method is a valid method for the assessment of the quality of the speech transmitted over communications channels intended for use with the Polish language. New STI curves were generated for the Polish speech intelligibility.

Key words: speech quality, speech intelligibility, logatom intelligibility, speech transmission index, STI.

1. Introduction

Speech intelligibility is the most important parameter of speech quality in analog and digital telecommunication channels, auditory rooms, public-address systems and for the selection of hearing aids [2, 4, 7, 10–13, 17, 18, 20, 21, 23, 25, 27, 28, 33]. The subjective intelligibility measure might be based on phonemes, logatoms, words and sentences [4, 6, 11, 17, 18, 28, 33]. The subjective measurement of speech intelligibility is very time-consuming, expensive and requires a large and a specially trained group of listeners.

Designers of the speech transmission equipment and systems lean towards objective measuring techniques often disregarding the application limitations and the measuring accuracy dependent on the type of the tested object and measuring conditions [1, 3, 5, 6, 8–11, 13–16, 19, 22–24, 26, 29–32, 35]. But the final verifier of the quality of the speech transmission equipment is its user – a human being – and the verification is made by subjective measurements of the transmission quality.

The quality of the transmission depends on the objective, the physical parameters of the devices used in transmitting speech signals and on subjective factors connected with the sender and receiver of the information transmitted. Subjective test results should be dependent mostly on physical parameters of the tested transmission channel, but not on the structure of the language test. At the semantic level, the information is eliminated by means of logatom sets used for the determination of logatom or phonemic articulation. Thus logatom articulation averaged over a set of listeners, functioning as a verifier of the effectiveness of the objective methods, is used as a reference measure of the transmission channel quality.

2. Measurement of logatom intelligibility

Formal subjective tests are described by the Polish Standard PN 90/T-05100 "Analog Telephone Chains. Requirements and Methods of Measuring Logatom Articulation" [33].

The subjective measurement of logatom 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 in the midst of persons who have normal hearing and normal experience in pronunciation in the language used in the test. A person is considered to have normal hearing if her/his auditory threshold does not exceed 10 dB for any frequency in the band of 125 Hz – 4000 Hz and 15 dB in the band of 4000 Hz – 6000 Hz. The hearing should be tested by means of a diagnostic audiometer. The size of the listening group should be such that the averaged test results obtained do not change as the group size is increased further (minimum 5 persons). The group of listeners who 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 write the received logatom at leisure. It is recommended that logatoms should be spoken with 3–5 s. pauses in between. The time interval between sessions should be not shorter than 24 hours 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 [33]. The notation should be legible to prevent a wrong interpretation of the logatom. The received logatoms may be written in a phonetic transcription (a group of specially trained listeners is needed for this) or in an orthographic form specific for the given language. A member of the team who checks listening test lists calculates, for each listener and for each logatom list, the number of correctly received logatoms and then determines the average logatom articulation.

3. Modulation transfer function

HOUTGAST and STEENEKEN described an objective method for estimating the speech intelligibility in rooms by calculating a physical index from the modulation transfer function (MTF) [8, 9, 19–22]. This physical index, called the speech transmission index (STI) is calculated at discrete frequencies, weighted, summed and normalized to yield a single index of speech intelligibility. This method was modified by BRACHMAŃSKI for Polish speech transmitted via analog telecommunication channels [3].

The test signal in the modulation transfer function is noise with a spectral shaping that fits the average speech spectrum. This noise is amplitude modulated in octave bands (7 octave bands) by a low frequency sinusoidal signal with a depth of modulation $m = 100\%$. The modulation frequency is chosen from 0.63 to 12.5 Hz according to the centre frequencies of the one-third octave bands covering this range. The depth of modulation m will be reduced due to room reflection, background noise, band pass limiting, etc. The depth of modulation at the receiver position, which is converted into an attenuation (in dB), is referred to as Transmission Index (TI). The Modulation Transfer Index (MTI) is given by the Transmission Index as a function of the modulation frequency. The MTI is determined for each octave band. Finally, the one single index called Speech Transmission Index (STI) is obtained by a weighted summation of the Modulation Transfer Index (MTI).

In 1985, the foreshortened adaptation of the Speech Transmission Index (STI) was developed by Houtgast and Steeneken and termed RASTI (Rapid Speech Transmission Index) [29, 30]. The RASTI method is restricted to the 500 Hz and 2000 Hz octave bands and to 4 and 5 different modulation frequencies, respectively. A handheld measurement instrument was engineered by Brüel & Kjaer. The RASTI system consists of a transmitter (Type 4225), which is placed at the speaker's position and a receiver (Type 4419) placed at the listener's position [34].

Another simplified version of STI is the STITEL (Speech Transmission Index for Telecommunication Systems) [29, 30, 35]. The STITEL method is applied at the same octave bands as the STI, but in each band only one modulation frequency is used. The test signal includes all seven octave bands and all of them are analysed simultaneously.

An efficient form of the Speech Transmission Index method for public address systems is STI-PA (Speech Transmission Index-Public Address) [29] The research revealed that the original concept may lead to some errors for conditions at which a non-contiguous transfer is used. HOUTGAST and STEENEKEN modified the original STI model [20–22]. The revised Speech Transmission Index (STI_r) is obtained by a weighted summation of the modulation transfer indices for all octave bands and the corresponding redundancy corrections. The redundancy correction is related to the contribution of adjacent frequency bands [29, 30].

4. Experiment

4.1. Measurement system

Objective and subjective measurements were performed in frequency bands of 100–6000 Hz and 400–2500 Hz and in the telephone band, i.e. 300–3400 Hz, for a signal/noise ratio ranging from -15 dB to $+15$ dB [2–5].

Different transmission conditions were simulated using a telecommunication channel MKT-1 model made in the Institute of Telecommunications and Acoustics of the Wrocław University of Technology.

In the model the following interference and distortions can be introduced into the transmitted signal:

1) additive interference:
   a) white noise the level of which is adjusted continuously,
   b) pink noise the level of which is adjusted continuously,
   c) hum the level of which is adjusted continuously,
   d) impulse interference with four different average values of the random distribution generating a time interval between successive pulses the amplitude of which is adjusted continuously;
2) frequency band limitation – 78 combinations set stepwise in a band of 100–4500 Hz and the so-called full band position; the limit frequency of the high-pass filter can be varied from 100 Hz to 500 Hz in steps of 100 Hz, whereas that of the low-pass filter can be changed from 2000 Hz to 4500 Hz in steps of 500 Hz; the full-band position represents a bandwidth of 100–6000 Hz;
3) attenuation-diagram linear distortions (9 switched independently on linear equalizers);
4) external interfering signals.

In the presented experiment the influences of: the white noise in the 100–6000 Hz band, the 400–2500 Hz band and the 300–3400 Hz band, the pink noise in the 100–6000 Hz band and the 400–2500 Hz band, the hum in the 400–2500 Hz band and the intelligible crosstalk (for Polish and English) in the 400–2500 Hz on speech intelligibility were studied. The crosstalk was produced by recording a text in Polish and English read by a speaker and then reproducing it during the articulation measurements. The reproduced text was fed into the telecommunication channel MKT-1 model through the input assigned for external interfering signals. The speech and interfering signal levels were controlled by means of a 2606 Bruel & Kjaer measuring instrument [2].

4.2. Subjective tests of logatom intelligibility

Subjective tests of logatom intelligibility were done at the Institute of Telecommunications and Acoustics at the Wrocław University of Technology according to the Polish Standard PN90/T-05100. A listening team was selected from the Wrocław University of Technology students with normal hearing. The qualification was based on audiometric tests. There were 12 persons in the team. The subjects were trained for two sessions (6-hour training). Four phonetically and structurally balanced 100-logatom lists per measuring point were used. The list was selected at random. Subjective measurements of logatom intelligibility were taken at conditions of monaural listening using electrodynamic receivers for the speech signal level of 80 dB(A) [2, 33].

4.3. STI based measurements

Measurements of the speech transmission quality were performed with the STI index meter designed at the Institute of Telecommunications and Acoustics of the Wrocław University of Technology (the MTF method) and the RASTI measurement system of Brüel & Kjaer (type 3361).

The testing signal was fed to the telecommunication channel MKT1 designed at the Institute of Telecommunications and Acoustics of the Wrocław University of Technology (the channel simulated various transmission conditions given in Sec. 4.1 – white and pink noises, frequency band and intelligible and unintelligible crosstalk). Then the signal from the output of the MKT1 channel is sent to the receiver side of the measurement system.

5. Discussion

Figure 1 presents the results obtained from the experiments: the characteristics determined by HOUTGAST and STEENEKEN [8, 9, 22] for the Dutch language, BRADLEY

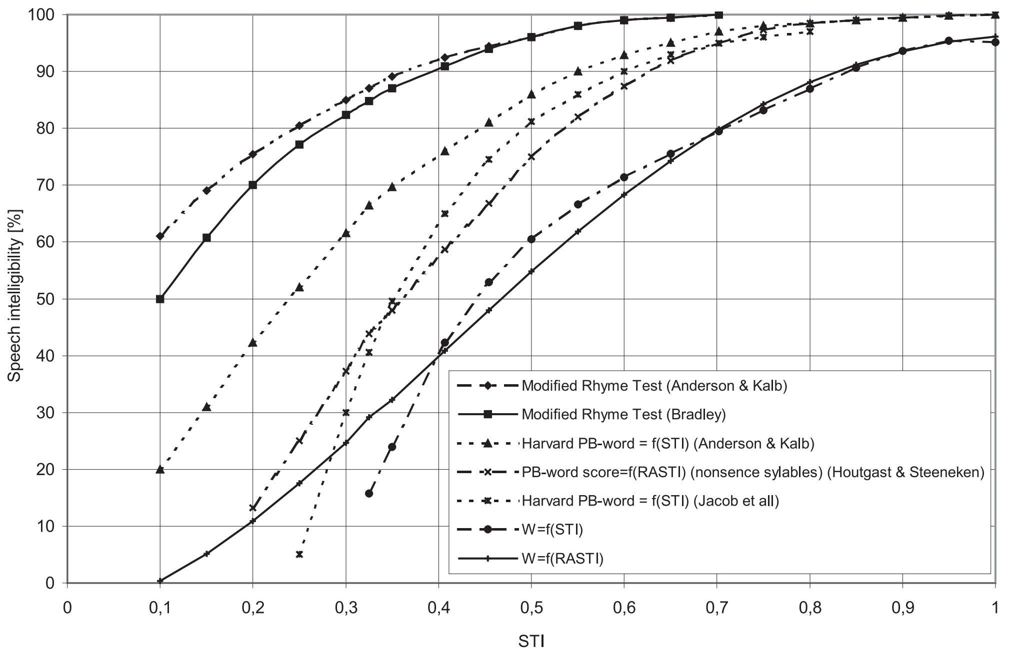

Fig. 1. Relationship between speech intelligibility and STI for different tests (the modified rhyme curve is estimated from the Kryter an Whitman data [1], the modified rhyme curve is derived from the BRADLEY data [6], the PB-word curve is derived from ANDERSON and KALB [1], the CVC nonsense syllable curve is redrawn from STEENEKEN and HOUTGAST [19]; the PB-word curve is derived from JACOB, BIRKLE and ICKLER [10], logatom curves ($W = f$(STI) and the $W - f$(RASTI) are derived from the data of the present work).

[6], JACOB, BIRKLE and ICKLER [10] and ANDERSON and KALB [1] for English and characteristics obtained for measurements with the Brüel & Kjaer system of type 3361 (RASTI) and with the MTF method (full STI) for Polish. The relation between the objective STI and subjective logatom intelligibility (W_L) for the Polish language was studied for 80 different transmission conditions. A fourth-order polynomial regression function was computed for speech intelligibility versus STI and RASTI. The regression function was obtained basing on the results

1) from the full STI method

$$W_L = -1447\mathrm{STI}^4 + 4221.74\mathrm{STI}^3 - 4609.47\mathrm{STI}^2 + 2308.57\mathrm{STI} - 378.68;$$

2) from the RASTI method

$$W_L = 141.46\mathrm{STI}^4 - 457.15\mathrm{STI}^3 + 402.18\mathrm{STI}^2 + 14.28\mathrm{STI} - 4.6083.$$

The analysis of the results obtained reveals a monotonous relation between the logatom intelligibility (W_L) and the speech transmission index (STI). The shape of the characteristic obtained from the measurements made with the RASTI system is similar to the characteristic representing the relation between logatom intelligibility and the STI index, especially for STI ≥ 0.4. It should be noticed that the differences between the two above curves are minimal and are due to the different measurement techniques used. It can be therefore concluded that the measurments done with method modified for the Polish language and the RASTI one give similar results. However, the differences between both the methods related to the Polish language and the characteristic given by Houtgast and Steeneken for Dutch are significant. Comparative studies performed by ADERSON and KLAB [1] for word tests used in the United States (Harvard PB-word test) revealed that also for the English language the verification of the results obtained by the MTF method is necessary. BRADLEY [6] measured intelligibility using the Fairbanks rhyme test. JACOB, BIRKLE and ICKLER [10] used phonetically balanced (PB) monosyllabic English words as specified by ANSI [25].

6. Conclusions

The results of the comparative studies show that the relationship between the STI and the speech intelligibility dependends strongly on the language and the type of test lists used. Figure 1 shows that subjects score much higher on the two rhyme tests than in other subject tests. This was expected because rhyme tests are multiple-choice tests. In addition, in the rhyme tests words differ only in one consonant. The test used by Houtgast & Steeneken was shown to be similar to the ANSI method used by Jacob, Birkle and Ickler.

Summing up, the RASTI method can be used for measures of the Polish speech transmission quality at the condition that there will be done the transformation of results obtained with the RASTI method (B & K measure system type 3361) related to Dutch into the scale obtained for Polish. It means that the RASTI result should be converted into the logatom intelligibility using relation (2). Next, the class of the telecommunication channel can be determined according to the table included in the recommendation number PN-90/T-05100.

In the case of using the RASTI method, the specific character of the Polish language and the type of the applied test have to be respected.

References

[1] ANDERSON W.B., KALB J.T., *English verification of STI method for estimating speech intelligibility of a communications channel*, JASA, **81**, 6, 1982–1985 (1987).

[2] BRACHMAŃSKI S., *Effect of additive interference on speech transmission*, Archives of Acoustics, **27**, 2, 95–108 (2002).

[3] BRACHMAŃSKI S., *Modulation Transfer Function (MTF) as measure of Polish speech transmission quality* [in Polish], Doctoral Dissertation, Wrocław 1982.

[4] BRACHMAŃSKI S., *Subjective measurement of speech transmission quality in telecommunication chanels* [in Polish], Proceedings of 46-th Open Seminar on Acoustics, Kraków–Zakopane, 99–104, 1999.

[5] BRACHMAŃSKI S., ZUK J., *Application of RASTI system for quality evaluation of Polish speech transmission in analog communication system* [in Polish], Proceedings of 44-th Open Seminar on Acoustics OSA'97, Gdańsk, 177–182, 1997.

[6] BRADLEY J., *Predictors of speech intelligibility in rooms*, JASA, **80**, 3, 837–845, 1986.

[7] FARINA A., *Acoustic quality of theatres: correlations between experimental measures and subjective evaluations*, Applied Acoustic, **62**, 889–916 (2001).

[8] HOUTGAST T., STEENEKEN H.J.M., *The Modulation Transfer Function in room acoustics as a predictor of speech intelligibility*, Acustica, **28**, 66–73 (1973).

[9] HOUTGAST T., STEENEKEN H.J.M., *Predicting speech intelligibility in rooms from the Modulation Transfer Function. I. General room acoustics*, Acustica, **46**, 60–72 (1980).

[10] JACOB K.D., BIRKLE T.K., ICKLER C.B., *Accurate prediction of speech intelligibility without the use of in-room measurements*, J. Audio Eng. Soc., **39**, 4, 232–242 (1991).

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

[12] MAJEWSKI W., MYŚLECKI W., BAŚCIUK K., BRACHMAŃSKI S., *Application of modified logatom intelligibility test in telecommunications, audiometry and room acoustics*, Proceedings 9-th Mediterranean Electrotechnical Conference Melecon'98, Tel-Aviv, Israel, 25–28, 1998.

[13] MAJEWSKI W., MYŚLECKI W., BRACHMAŃSKI S., *Methods of assessing the quality of speech transmission*, 47-th Open Seminar on Acoustics, Rzeszow, 66–75, 2000.

[14] MAJEWSKI W., ZALEWSKI J., *Articulation index of polish speech* [in Polish], Doctoral Dissertation, Wrocław 1964.

[15] MAPP P., *A comparison between STI and RASTI Speech Intelligibility Measurement Systems*, The 111-th AES Convention, Los Angeles, Preprint 5668, 2002.

[16] MAPP P., *Limitations of current sound system intelligibility verification techniques*, The 111-th AES Convention, Los Angeles, Preprint 5668, 2002.

[17] SOTSCHEK J., *Methoden zur Messung der Sprachgüte I: Verfahren zur Bestimmung der Satz- und der Wortverständlichkeit*, Der Fernmelde Ingenieur, **10**, 1–31 (1976).

[18] SOTSCHEK J., *Sprachverständlichkeit bei additiven Störungen*, Acustica, **57**, 257–267 (1985).

[19] STEENEKEN H.J.M., HOUTGAST T., *A physical method for measuring speech-transmission quality*, JASA, **67**, 1, 318–326 (1980).

[20] STEENEKEN H.J.M., HOUTGAST T., *Mutual dependence of the octave-band weights in predicting speech intelligibility*, Speech Communication, **28**, 109–123 (1999).

[21] STEENEKEN H.J.M., HOUTGAST T., *Phoneme-group specific octave-band weights in predicting speech intelligibility*, Speech Communication, **38**, 399–411 (2002).

[22] STEENEKEN H.J.M., HOUTGAST T., *Validation of revised STI method*, Speech Communication, **38**, 413–425 (2002).

[23] TISSEYRE A., MOULINIER A., ROUARD Y., *Intelligibility in various rooms: Comparing its assessment by (RA) STI measurement with a direct measurement procedure*, Applied Acoustic, **53**, 1–3, 179–191 (1998).

[24] WIJNGAARDEN S.J., STEENEKEN H.J.M., *Objective prediction of speech intelligibility at high ambient noise level using the speech transmission index*, Proc. of the 6-th Europeen Conf. on Speech Comm. and Technol., Eurospeech 99, Budapest, **6**, 2639–2642, 1999.

[25] ANSI S 3.2, *Methods for measuring the intelligibility of speech over communication systems*, 1989.

[26] ANSI S 3.5, *Methods for the calculation of the speech intelligibility index (SII)*, 1997.

[27] CCITT, *Quality of telephone transmission, local networks and telephone sets* [in Polish], CCITT's Green Book, Vol. V, WKiŁ, Warsaw 1976.

[28] ISO/TR 4870, *Acoustics – The construction and calibration of speech intelligibility tests*, 1991.

[29] IEC 60268-16, *Sound system equipment – Part16: Objective rating of speech intelligibility by speech transmission index*, Ed. 3, International Electrotechnical Commission, 2003.

[30] ITU-T Recommendation P.861, *Objective quality measurement of telephone-band (300-3400 Hz) speech codecs*, 1996.

[31] ITU-T Recommendation P.862, *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs*, 2003.

[32] Polish Standard PN-EN-60268-16, *Sound system equipment – Part16: Objective rating of speech intelligibility by speech transmission index*, 2002.

[33] Polish Standard PN-90/T-05100, *Analog telephone chains. requirements and methods of measuring logatom articulation* [in Polish], 1990.

[34] *Speech transmission meter type 3361, instruction manual*, Brüel & Kjaer.

[35] *Measuring speech intelligibility using DIRAC*, Technical Note TN002, Brüel & Kjaer, 2003.