Acoustic and Biomedical Engineering

Speech Intelligibility Measurements in Auditorium

K. Leo*

Faculty of Physics and Applied Mathematics, Technical University of Gdańsk

G. Narutowicza 27/12, 80-233 Gdańsk, Poland

Speech intelligibility was measured in Auditorium Novum on Technical University of Gdansk (seating capacity 408, volume 3300 m^3). Articulation tests were conducted; STI and Early Decay Time EDT coefficients were measured. Negative noise contribution to speech intelligibility was taken into account. Subjective measurements and objective tests reveal high speech intelligibility at most seats in auditorium. Correlation was found between spatial differences in responses to articulation tests and EDT for low frequencies divided by EDT for high frequencies.

PACS numbers: 43.55 Hy

1. Introduction

1.1. Acoustical parameters in auditorium

During the process of building Auditorium Novum at Technical University of Gdańsk, concern of acoustical quality of the hall arouse. Taking the primary function of auditorium into account, high speech intelligibility was judged to be the most desirable parameter through the audience area in the hall [1].

To achieve high speech intelligibility in auditorium, a number of acoustic parameters have to be considered, for instance, reverberation time RT60 should have the optimum length, starting from minimum value 0.7 s for room of volume 283 m^3. Then it can linearly increase with the logarithm of room volume up to 1.1 s with deviations less than 10% [2]. This value prevents masking of speech sounds by late reflections [3]. On the other hand, it supports enough sound energy through the hall. Late reflections are not only important in speech intelligibility process but also much attention has to be paid to early reflections, as they play crucial role in the process of transporting intelligible speech from the speaker to listeners.

1.2. Mechanism of loss of speech intelligibility in relation to early reflections

The nature of speech sounds determines the mechanism of loss of speech intelligibility in enclosed space. Vowels and consonants convey different sound energy. The average level of consonants is 10–12 dB lower than the level of vowels. Thus, consonants can be easily masked by vowels in reverberant space. Moreover, the most energy of consonants is at higher frequencies, vowels are represented at lower and middle frequency bands [5]. Dividing audible spectrum into 10 equal intelligibility parts, one can see that frequencies up to around

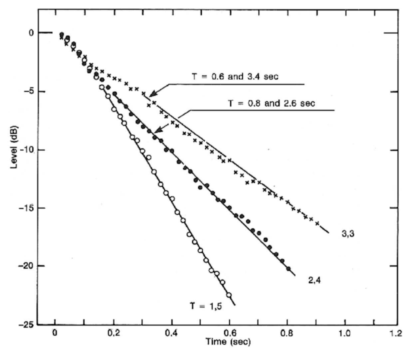

Fig. 1. Three sound decay curves for different rooms with different RT60 length. Although values of RT60 differ significantly, the rooms share the same shape of decay curve for early reflections. Speech intelligibility is equal in each of these rooms [4].

250 Hz and above 7.5 kHz do not add to intelligibility of speech [6]. Sounds of frequencies below 250 Hz act just as maskers for higher frequencies speech sounds.

Finally, when comparing average length of vowels and consonants, consonants are shorter than vowels, i.e. 20 ms compared to 90 ms. Vowels are stationary signals, with strong harmonic content, consonants are often non-stationary in character (e.g. plosives), rather without harmonic content, sometimes with strong noise content (e.g. fricatives).

To prevent the masking process of consonants by vowels, enough sound energy from the speaker to listeners has to be transported in a time range corresponding to average length of speech sound. Otherwise, the listener will hear elements of speech overlapping each other, which will deteriorate intelligibility. This critical time interval was found in the area of around 50 ms [5].

This is already known process of masking elements of speech sounds by other elements of speech. What should be moreover considered in a reverberant space? Is there an additional mechanism of deteriorating speech intelli-

* corresponding author; e-mail: krzysztof.leo@gmail.com

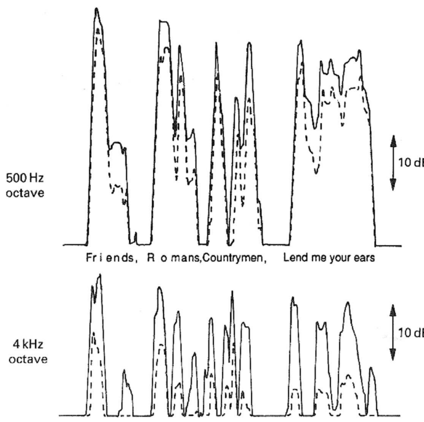

Fig. 2. Sound level vs. time for speech phrase at low mid and high frequencies. Solid line — in front of the speaker, dotted line — behind speaker [5].

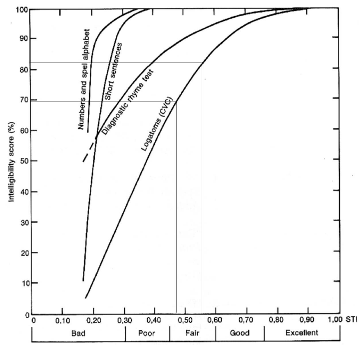

Fig. 3. Typical relations between the STI and intelligibility scores for various types of tests, among them is logatoms' test [7]. Range of STI values measured in a hall and corresponding intelligibility scores is marked.

gibility by early reflections?

When strong early reflections are decaying, they are interfering with direct sound. Speech signal becomes distorted by discrete, delayed signals. This process is likely dependent on the relative length of the speech sounds and decay of early reflections. High frequency speech sounds are more likely to be misunderstood in conditions of given EDT than lower speech sounds. When high frequency elements of speech are being distorted by early reflections, lower frequency sounds remain still distinctive to the listener. This is mainly due to the fact of non-stationary character of consonants comparing to the stationary vowels. Thus speech intelligibility is likely to be dependent on the shape of EDT curve. The slope of EDT in the direction of higher frequencies curve should correspond to the better speech intelligibility. Flat EDT curve should mean poorer intelligibility.

1.3. Speech Transmission Index (STI) in relation to the results of articulation tests

In order to check validity of articulation tests, additional STI measurements were performed (see point 3.2). Figure 3 compares various types of intelligibility tests in relation to the STI coefficient (data taken from the literature). These results were obtained for the hall with listeners, for many languages. On the Fig. 3 one can see marked values of the STI with corresponding with intelligibility scores. This area is in special interest of this work.

2. Method

2.1. Measurements

Articulation tests in which 101 participants (aged 22–24) took part, were carried. Listeners were asked to cover the whole seating area uniformly. Two syllabic lists, consisting of 100 nonsense elements (logatoms) were read. The lists were balanced for Polish language. First list was read without air-conditioning noise, the second one with fans being switched on.

Objective measurement system consisted of sound source — a gunshot, omni directional measurement microphone, sound pressure level meter and PC with sound card. Measurements were evaluated without audience. Eighteen measuring points were determined.

2.2. Analysis

Three impulse responses from each point were added in time domain and averaged output impulse response was analyzed. STI values were calculated and RT60 and EDT in octave bands from impulse responses were obtained. EDT was measured for early decay, from 0 dB to −10 dB (±3 dB).

3. Results

3.1. Articulation test

Articulation tests revealed high speech intelligibility in the hall. Average intelligibility is 88.5% in the absence of disturbing noise from air-conditioning system and 80.8% in presence of noise. Figures 4 and 5 present spatial differences of intelligibility. Numbers indicate Percentage Speech Articulation (PSA). Contours of equal speech intelligibility were drawn in order to show areas of better and poorer speech intelligibility. Averaging of 3 values of PSA in the radius of approximately 2 m from each measurement point was done (p1 to p18 indication on figures). Values written under measurement point are three PSA averaged.

Averaged PSA values in different seats in the hall are shown in the Table I. Differences of speech intelligibility through the hall were not very large. Maximum difference between average responses in best intelligibility rated place (p6) and worse place (p18) is 8.4% (noise OFF) and 9.7% (noise ON).

Level and spectrum of noise in the hall was measured in the absence of audience in the middle of the hall (p2). Figure 6 shows spectrum of background noise in

MOST WIEDZY Downloaded from mostwiedzy.pl

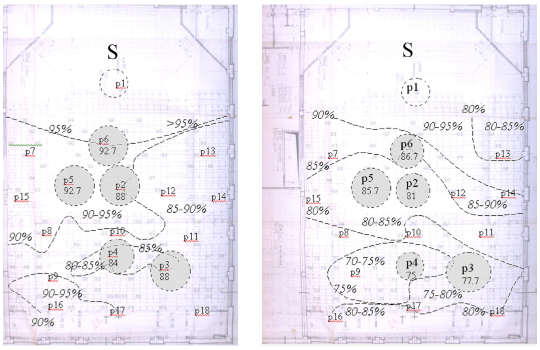

Fig. 4. left — noise OFF, Fig. 5 right — noise ON. Speech intelligibility equal contours and eighteen measurement points with averaged PSA around sixteen of them. Points on gray background are six survey points shown on EDT plots. Letter S marks position of the sound source.

the presence and absence of disturbing noise from air-conditioning system. There are large differences in level of noise at low frequencies.

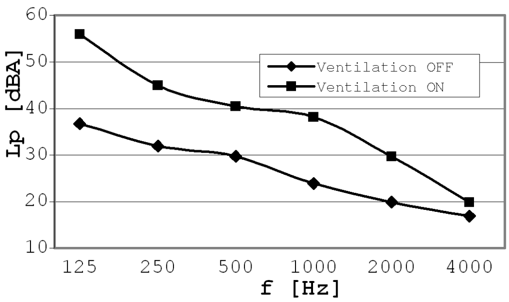

Fig. 6. Spectrum of background noise in presence and absence of air-conditioning noise.

The total level of noise with air-conditioning off was: 33.9 dBA, 30.1 dBA, 30.3 dBA at points 1, 2, 3 and with air-conditioning on was: 45.4 dBA, 44.8 dBA 43.7 dBA. There were not large differences between levels of background noise noted through the hall, apart from 2–4 dB rise in the front of the hall (p1). This is probably caused by part of air-conditioning system located over the stage always being switched on.

3.2. STI values

STI values calculated from impulse responses of the hall do not depend on the speech intelligibility scores measured in the hall. Table shows STI coefficients measured from averaged impulse responses in time domain to compare them with averaged PSA.

Not all values of STI correspond to values of averaged PSA in the way shown on the Fig. 3. STI is not enough precise to measure small differences of PSA. Region of STI marked by minimum and maximum values — 0.51 and 0.57 correlates with the area 73% — 88% of PSA in Fig. 3. However this region is not far below the average speech intelligibility measured by articulation test — 88.5%.

TABLE I

STI values and corresponding average Percentage Speech Articulation, NOISE OFF.

Measuring point	STI	Averaged PSA [%]
1	0.54	–
2	0.52	88
3	0.56	88
4	0.57	84
5	0.55	92.7
6	0.51	92.7
7	0.6	90
8	0.6	91
9	0.58	91.3
10	0.58	88.7
11	0.57	87
12	0.55	88
13	0.56	89.3
14	0.55	–
15	0.62	91.7
16	0.61	90.2
17	0.59	89
18	0.59	84.3

3.3. RT60 curves

RT60 values do not vary significantly through hall. Figure 7 shows RT60 from six measurement points in octave bands. RT60 value shows more discrepancies in lower bands (63 Hz and 125 Hz) bands than in higher bands.

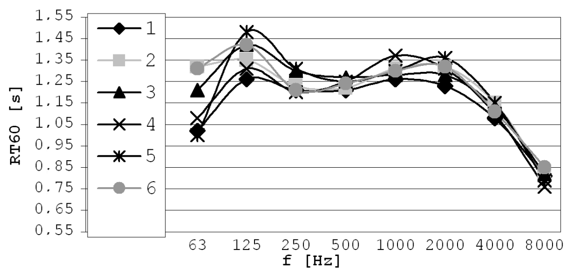

Fig. 7. RT60 curves for 6 measurement points. Points indicated with gray background and indicators p1, p2, p3, p4, p5 and p6 on the Fig. 4.

3.4. EDT curves

EDT values for octave bands were obtained in order to find better range of differences in that parameter. Values

MOST WIEDZY Downloaded from mostwiedzy.pl

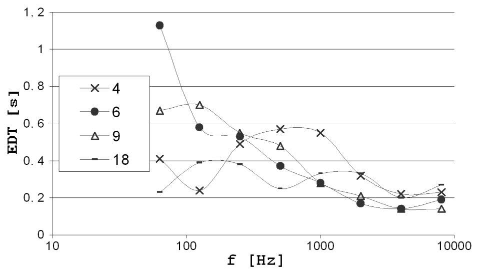

Fig. 8. EDT values for 4 measurement points. Points indicated with indicators p4, p6, p9, p16 on the Fig. 4.

of EDT in octave bands show differences, when different measuring points are compared.

Figure 8 shows that in seats with poorer intelligibility (point 4, 18) EDT curves are falling slower to the direction of higher frequencies. In points with better intelligibility (point 6, 9) EDT curves are falling more rapidly towards higher frequencies.

To measure the ratio of EDT slope towards high frequencies, a coefficient EDT_s (for EDT slope) was proposed. (EDT_s = EDT for 63 Hz plus EDT for 125 Hz divided by EDT for 4 kHz plus EDT for 8 kHz). Values of EDT_s with corresponding averaged PSA are shown in Table II.

TABLE II

EDT_s coefficient with corresponding, averaged PSA in 16 measurement points in the presence and absence of noise.

Measuring point	Averaged PSA [%] noise ON	Averaged PSA [%] noise OFF	EDT_s
1	–	–	2.25
2	81	88	2.47
3	77.7	87.7	2.33
4	75	84.3	1.44
5	85.7	92.7	3.4
6	86.7	92.7	5.2
7	86	90	4.2
8	80	91	5.1
9	76.3	91.3	4.9
10	81	88.7	4
11	84	87	1.9
12	86	88	2.6
13	83.3	89.3	3.5
14	–	–	1.4
15	83.3	91.7	4.9
16	78.5	90.2	5.7
17	78.7	89	3.1
18	77	84.3	1.3

Values of coefficient EDT_s at sixteen points against 3 PSA averaged in the condition of presence of the noise are plotted in Fig. 9 and in the condition of absence of noise are plotted in Fig. 10.

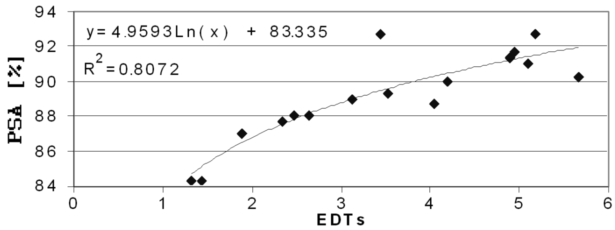

Fig. 9. EDT_s coefficient against corresponding, averaged PSA in sixteen measurement points for condition without disturbing noise.

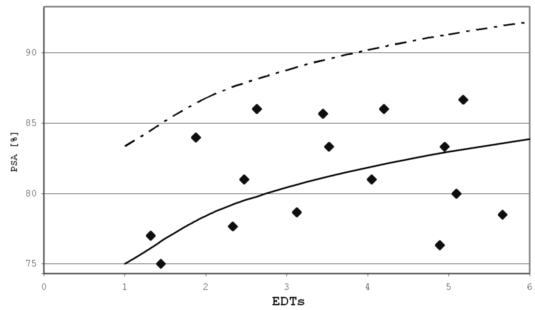

Fig. 10. EDT_s coefficient against corresponding, averaged PSA in 16 measurement points for condition with disturbing noise. Dotted line presents tendency obtained without background noise, solid line shows estimated tendency with background noise. Solid line was drawn using equation $y = 4.95 \ln(EDT_s)+75$ comparing to the dotted line using equation $y = 4.95 \ln(EDT_s) + 83.335$.

Correlation ($R^2 = 0.81$) has been found between the level of EDT_s and average PSA in the condition without noise. This fact suggests the dependence of Percentage Syllable Articulation on EDT curve slope. The mechanism of deteriorating speech intelligibility by early reflections, and dependence of intelligibility on the shape of EDT probably occurred in measured auditorium. Noise highly, and randomly deteriorated intelligibility in the hall. Although weak correlation was found in the presence of disturbing noise from ventilating fans, correlation between the EDT_s (ratio of EDT for low frequencies to the EDT for high frequencies) with speech intelligibility scores changed into observable rising tendency. Logarithmic equations used to model this phenomena show differences in values of their constants. In condition without noise logarithmic model describes phenomena much better than with background noise.

4. Conclusions

PSA articulation tests were conveyed in the auditorium of seating capacity 406 and volume 3300 m^3.

MOST WIEDZY Downloaded from mostwiedzy.pl

A map of different level of speech intelligibility was constructed. STI measurements showed no significant differences through the audience area, although the mean value of PSA corresponds to mean STI value. The need of more precise objective measure of speech intelligibility arouse. Logarithmic correlation ($R^2 = 0.81$) was found between EDT_s and averaged PSA. The model of deteriorating speech intelligibility in connection to different profiles of EDT curve versus frequency was proposed.

References

[1] M. Barron, *Proc. Institute of Acoustics*, **8** Part 3, 371 (1986).

[2] V.O. Knudsen, *Acoustical designing in Architecture*, Ch. 9, Acoustical Society of America, Los Angeles 1988.

[3] H. Kurtovič, *Acoustica* **33**, 32 (1975).

[4] H.J. Steeneken, T. Houtgast, in: *Rasti, Technical review*, Brüel & Kjær 1985, p. 13.

[5] M. Barron, *Auditorium Acoustics and Architectural Design*, Chapman & Hall/Routledge, London 1993, Ch. 6, 7.

[6] H.K. Dunn, S.D. White, *J. Acoust. Soc. Am.* **11**, 278 (1940).

[7] T. Houtgast, H.J. Steeneken, in: *Rasti, Technical review*, Brüel & Kjær 1985, p. 3.

MOST WIEDZY Downloaded from mostwiedzy.pl