Biological Description of Inner Ear using Telecommunications and Power Engineering

Violeta Stojanovic, Zoran Milivojevic, Zoran Veliėkovic
College of Applied Technical Sciences of Niš, 20
Aleksandra Medvedeva, St, 18000 Niš, Serbia
{violeta.stojanovic@vtsnis.edu.rs}

ABSTRACT: *A broader perspective involving medicine and technology is treated in this paper. For solving biological issues, we evolve a fluid model by using both power engineering and signal processing. This approach brings real life solution to the issues in a wider sense. We presented a biological description of inner ear, simulation model for cochlea, pressure contours on the basilar membrane. This research shows how the interdisciplinary approach can solve real life issues.*

Keywords: Room Impulse Response (RIR), STI, MOS Test, Bablle Noise, Subjective Speech Intelligibility

Received: 4 September 2020, Revised 7 December 2020, Accepted 13 December 2020

DOI: 10.6025/stj/2021/10/1/7-14

1. Introduction

Verbal communication is a process of deliberate and, above all, understandable exchange of notions, thoughts and messages between people. For the efficient verbal communication, it is essential to provide good transmission of the speech signal from a speaker to a listener. Noises are one of the factors which affect speech intelligibility. Babble Noise (BN) is on of the best noise for covering the speech. [1]. It represents the acoustic nuisance which originates from N subjects speaking simultaneously end degrades the useful acoustic signal. Degradation measure is shown by SNR relation.

French, Steinberg and Beranek 1947 were the first who highlight the problem of transmission and speech intelligibility [2]. Houtgast and Steeneken in 1980 confirmed an objective method to measure the quality of voice transmission in the room [2] and suggested the acoustic parameters of the Speech Transmission Index, STI [3]. Subjective systems evaluation for transmitting voice using open and closed tests intelligibility, were first performed by Fletcher and Steinberg in 1929 [2], then Egan in 1944, Miller and Nicely in 1955, House in 1965 and Voiers in 1977 etc. Review of tests for evaluating intelligibility of speech was given by Pols in 1991 and Steeneken in 1992.

This paper estimates the Babble Noise impact on subjective speech intelligibility for “Octagon” room, which is situated at “Queen Mary” University in London. The experiment is consisted of two stages: a) the analysis of subjective speech intelligibility in a room under the influence of reverberation. [4] and b) the analysis of subjective speech intelligibility in a room under the influence of Babble noise. This experiment isbased on MOS test of speech signal intelligibility from a base

Serbian Matrix Sentence Test, SMST [5]. In the first stage of this experiment, the test group of listeners estimated the speech signal intelligibility obtained using the convolution of a clear speech signal out of SMST base and recorded impulse response. The impulse responses refer to the "Octagon" room at "Queen Mary" University in London and they were recorded by Centre for Digital Music [6] and they are available at http://c4dm.eecs.qmul.ac.uk/rdr/handle/123456789/10. In the second stage of the experiment, the test group estimated simulated acoustic signal intelligibility of the room with superimposed BN for predefined SNR relations. Two test groups of listeners participated in the experiment. After the analysis of the obtained results from the applied MOS tests, the conclusion about the BN impact on the subjective speech intelligibility in the room.

Work organization is as follows: Section 2 describes the experiment and the results shown with charts and graphs, Section 3 shows the results analysis and the Conclusion is in the Section 4.

2. Experiments

This paper describes the experiment which estimates BN impact on the subjective speech intelligibility in "Octagon at the Mile End campus of Queen Mary" at the University in London.

The objective acoustic parameter STI was determined using the impulse of the room in 25 measuring points MP (out of 169) whose positions are shown in Figure 1. [6] The analysis was carried out using software packages EASERA and Matlab.

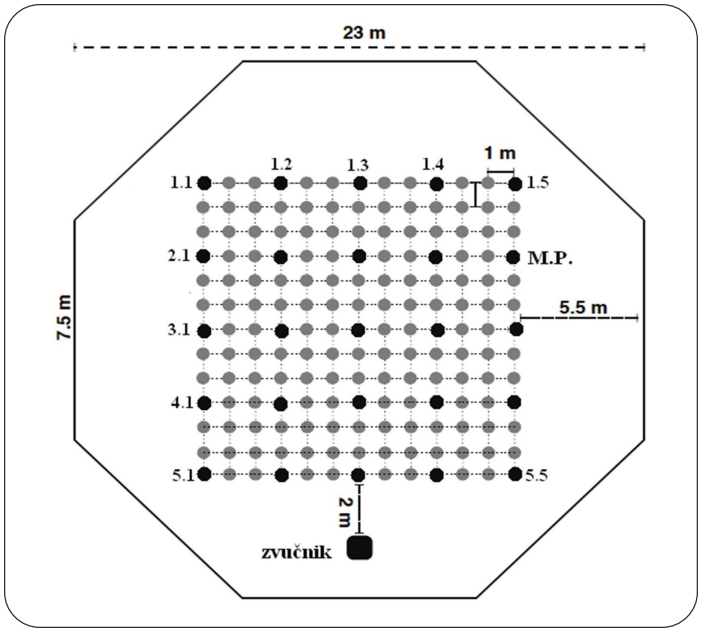

Figure 1. The position of the measuring points in the room

The subjective speech intelligibility in the room under the reverberation, [4], as well as the subjective speech intelligibility under the influence of BN were obtained using MOS test.

The subjective speech intelligibility was carried out for the MP (3, 4). This measuring point was chosen because the value of its parameters STI was the median of all the values within the set of measuring points. MP (3, 4) is 8.54 m away from the sound source.

In Figure 2 shows the model of the second applied stage of the experiment where: x - represents clear speech signal, h - is impulse response of the room, y - is generated acoustic signal which represents an equivalent to the speech signal recorded in that room, BN8 – BN of 8 speakers, k - coefficient for determining SNR, and z - generated acoustic signal representing simulated acoustic signal of the room with superimposed BN for predefined SNR relations.

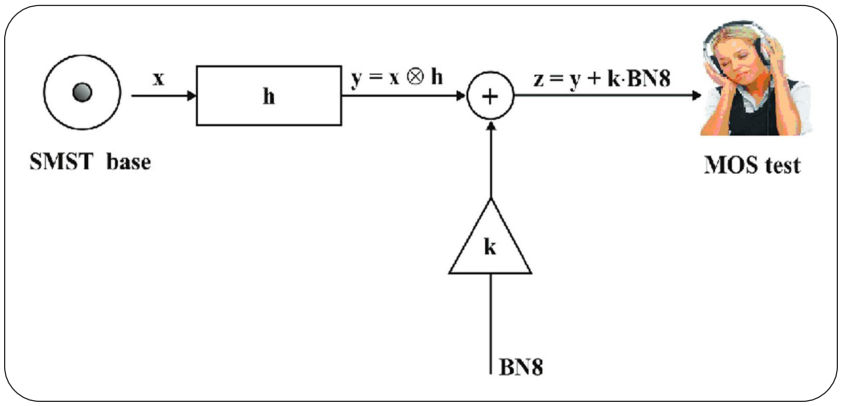

Figure 2. Experiment model

During MOS testing, a listener is in the acoustically isolated room and listens to the speech signal with earphones. After the emission of the sentence, which is semantically unexpected for both listener and examiner, the listener loudly repeats the sentence the way they understood it. The examiner records the results of correctly repeated sentence or correctly repeated certain PB words of the sentence. Each listener of the test group, consisted of 20 listeners, listens to 10 sentences in first stage of the experiment [4], and 6 sentences, consisted of 5 PB words, in the second stage of the experiment.

Based on the statistic analysis of MOS tests results, it was determined for male listeners, female listeners and entire test group in the room: a) under reverberation influence, the subjective intelligibility of correctly repeated whole sentences $(SI_{WS})_r$ or particular PB words $(SI_{PBw})_r$ and b) under BN influence, the subjective intelligibility of correctly repeated whole sentences $(SI_{WS})_{BN}$ or particular PB words $(SI_{PBw})_{BN}$ The estimation of **BN** influence on the subjective speech intelligibility in the room was defined by:

$$(\Delta SI_{WS})_{BN(SNR)} = (SI_{WS})_r - (SI_{WS})_{BN(SNR)}\,(\%) \tag{1}$$

$$(\Delta SI_{PBw})_{BN(SNR)} = (SI_{PBw})_r - (SI_{PBw})_{BN(SNR)}\,(\%) \tag{2}$$

where: $(\Delta SI_{WS})_{BN(SNR)}$ - the estimation of BN influence (for certain SNR value) on the subjective intelligibility of the entire sentences and $(\Delta SI_{PBw})_{BN(SNR)}$ - the estimation of BN influence (for certain SNR value) on the subjective intelligibility of the PB words. SNR = {20, 10, 5, 0, -1, -2} dB.

2.1 The Basis

To do the experiment, three bases were formed: a) the speech signal base, b) the impulse responses base and c) BN base.

a) The speech signal base end SMST [5]. SMST is consisted of isolated PB words pronounced in Serbian, divided into Proper Nouns (PN), verbs (V), numbers (N), adjective (A) and common noun (CN). According to program and a random law there can be formed, one of the possible a 100000 combinations, of sentences formatted: PN, V, N, A and CN. SMST base is phonetically balanced. Sampling frequency is f_s = 44.1 kHz with 2 BpS.

b) The impulse responses base contains audio recordings of “Octagon”, room impulse responses, recorded by “Centre for Digital Music” from “Queen Mary” University in London. [6]. “Octagon” is consisted of 8 walls 7.5 m high, which are covered with books, with floor and ceiling coated with plaster. The ceiling dome is 21 m high. The volume is 9500 m3 . Reverberation time is: RT_{30} (500 Hz) = 2.99 s end RT_{30} (1000 Hz) = 3.23 s.

The measurements of impulse responses were carried out using incentive log - sweep signal for 2 s [6] with f_s = 96 kHz, with 3 BpS. During the experiment, “Genelec” 8250A loudspeaker and omnidirectional DPA 4006 microphone were used.

c) BN was achieved by 8 speakers, 4 male and 4 female (**BN8**). Signal simulation was carried out SNR = {20, 10, 5, 0, -1, -2} dB.

2.2 Test Group

Test group is consisted of the students from The College of Applied Technical Sciences in Nis. The structure of test group is 10 male and 10 female students listeners, aged 19 ÷ 25. Mean value age is $\mu = 21.45$ yr. with standard deviation $\sigma = 2.16$ yr.

2.3 The Results

In Tables 1 – 3 shown in values $(SI_{WS})_{BN}$ end $(SI_{PBw})_{BN}$ by female and male listeners and the whole group with presence of BN8 for defining of SNR values, respectively. For the same group of listeners, the table 4 shows the value of influence assessment **BN8** (SNR), the subjective intelligibility entire sentences ΔSI_{WS} and PB words ΔSI_{PBw}. Figure 3. show the dependence: a) $SI_{WS} = f(SNR)$ end b) $SI_{PBw} = f(SNR)$ with presence of **BN8** by both male and female listeners. Figure 4. show the dependence $SI_{WS} = f(SNR)$ and $SI_{PBw} = f(SNR)$ with presence of BN8 for the entire whole group. Figures 5. and 7. shows the values $(SI_{PBw})r$ and $(SI_{PBw})_{BN(SNR)}$ by female and male listeners, and the whole test group, respectively. For the entire test group, in Fig. 6. can see value $(SI_{WS})r$. and $(SI_{WS})_{BN}$. In Figures 5. – 7. are marked with C_1 to the value of speech intelligibility group of listeners, or the whole of the test group in the room as a result effected by the reverberation. $C_2 \div C_7$ markings are assigned a value of speech intelligibility group of listeners, or the whole of the test group in the room due to **BN8** effected by *SNR* = {20, 10, 5, 0, -1, -2} dB.

SNR (dB)		**20**		**10**		**5**		**0**		**-1**		**-2**	
No. WS		3		2		0		0		0		0	
No. PBw	***PN***	7	33	6	24	4	19	2	5	2	3	0	0
	V	7		5		6		0		0		0	
	N	7		3		5		2		1		0	
	A	6		5		2		1		0		0	
	CN	6		5		2		0		0		0	
$(SI_{WS})_{BN}$ ***(%)***		30		20		0		0		0		0	
$(SI_{PBw})_{BN}$***(%)***		66		48		38		10		6		0	

Table 1. The subjective intelligibility of entire sentences and PB words by female listeners with presence of BN8 for defining of SNR values

SNR (dB)		**20**		**10**		**5**		**0**		**-1**		**-2**	
No. WS		3		3		1		0		0		0	
No. PBw	***PN***	8	31	4	29	3	19	5	11	0	4	3	6
	V	4		6		2		2		1		1	
	N	7		5		5		3		1		0	
	A	5		7		5		1		1		0	
	CN	7		5		4		0		1		2	
$(SI_{WS})_{BN}$ ***(%)***		30		30		10		0		0		0	
$(SI_{PBw})_{BN}$***(%)***		62		58		38		22		8		12	

Table 2. The subjective intelligibility of entire sentences and PB words by male listeners with presence of BN8 for defining of SNR values

3. The Results Analysis

Based on results shown in the Tables 1 – 4 and in the Figures 3. - 7. we can draw the following conclusions:

SNR (dB)		**20**		**10**		**5**		**0**		**-1**		**-2**	
No. WS		6		5		1		0		0		0	
No. PBw	*PN*	15	64	10	51	7	38	7	16	2	7	3	6
	V	11		11		8		2		1		1	
	N	14		8		10		5		2		0	
	A	11		12		7		2		1		0	
	CN	13		10		6		0		1		2	
$(SI_{WS})_{BN}$ *(%)*		30		25		5		0		0		0	
$(SI_{PBw})_{BN}$*(%)*		64		51		38		16		7		6	

Table 3. The subjective intelligibility of entire sentences and PB words by whole group with presence of BN8 for defining of SNR values

SNR (dB)	**Female listeners**		**Male listeners**		**The whole test group**	
	ΔSI_{WS} (%)	ΔSI_{PBw} (%)	ΔSI_{WS} (%)	ΔSI_{PBw} (%)	ΔSI_{WS} (%)	ΔSI_{PBw} (%)
20	22	19.6	28	21.2	25	20.4
10	32	37.6	28	25.2	30	33.4
5	52	47.6	48	45.2	50	46.4
0	52	75.6	58	61.2	55	68.4
-1	52	79.6	58	75.2	55	77.4
-2	52	85.6	58	71.2	55	76.4

Table 4. The value of influence assessment BN8 (SNR), the subjective intelligibility entire sentences and PB words for female, male and the whole group listeners

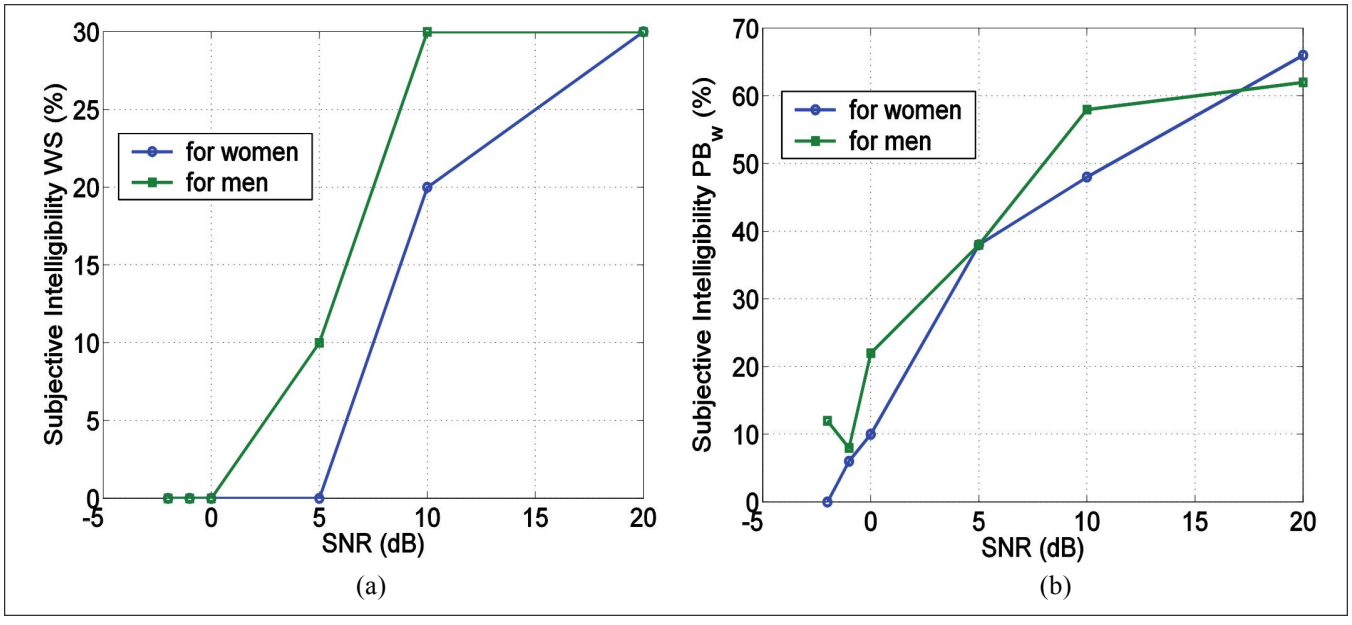

Figure 3. Dependence of: a) the entire sentences and b) the BP words, subjective intelligibility of SNR with presence of BN8 by both male and female listeners

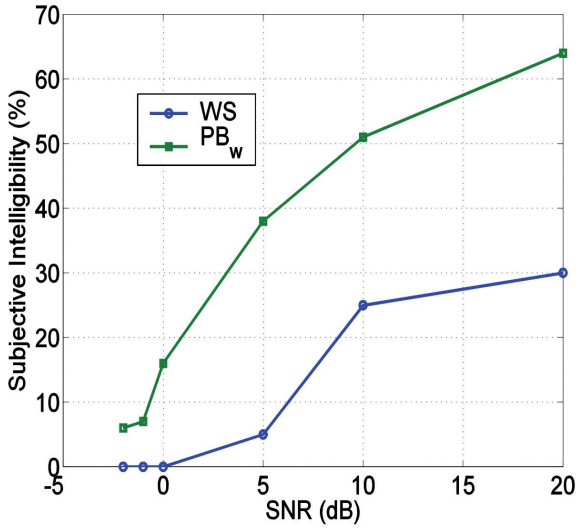

Figure 4. Dependence of the entire sentences end BP words subjective intelligibility of SNR with presence of BN8 for the entire whole group listeners

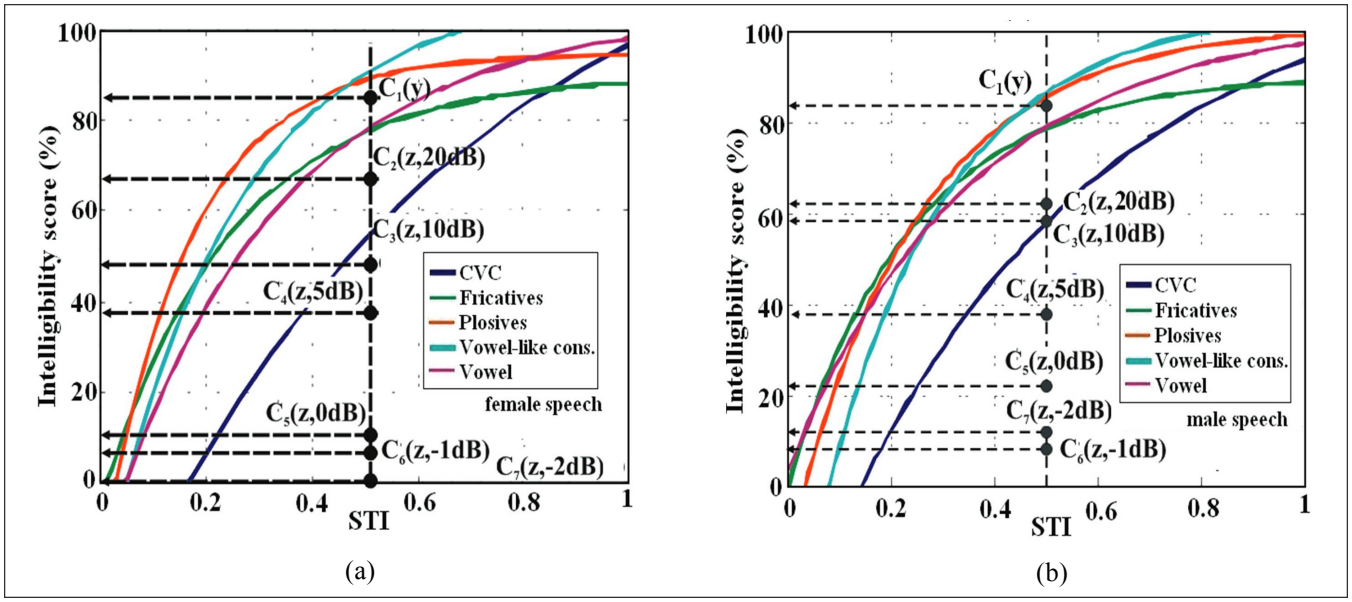

Figure 5. The subjective intelligibility of the PB words by: (a) female and (b) male, listeners in the room with presence of BN8

1) the subjective intelligibility of female listeners with the influence of BN8 is: a) the best on SNR = 20 dB, $(SI_{WS})_{BN}$ = 30 % and $(SI_{PBw})_{BN}$ = 66 % and b) 0% on SNR = {5, 0, –1, –2}dB and *SNR* = –2 dB for the whole sentences and BP words, respectively. The obtained values with the influence of reverberation are: $(SI_{WS})_r$ = 52 % and $(SI_{PBw})_r$ = 85.6 % [4]. Estimated values of the influence of BN8 on the speech intelligibility of female listeners are: $(\Delta SI)_{WS\,BN}$ = 22 % and $(\Delta SI_{PBw})_{BN}$ = 19.6%.

2) a) the subjective intelligibility of male listeners with the influence of **BN8** is $(SI_{WS})_{BN}$ = 30 % and $(SI_{PBw})_{BN}$ = 62 %) on SNR = 20 dB. The obtained values with the influence of reverberation are: $(SI_{WS})_r$ = 58 % and $(SI_{PBw})_r$ = 83.2 % [4]. The estimated values of the influence of BN on the speech intelligibility of male listeners are: $(\Delta SI_{WS})_{BN}$ = 28 % and $(\Delta SI_{PBw})_{BN}$ = 21.2%.

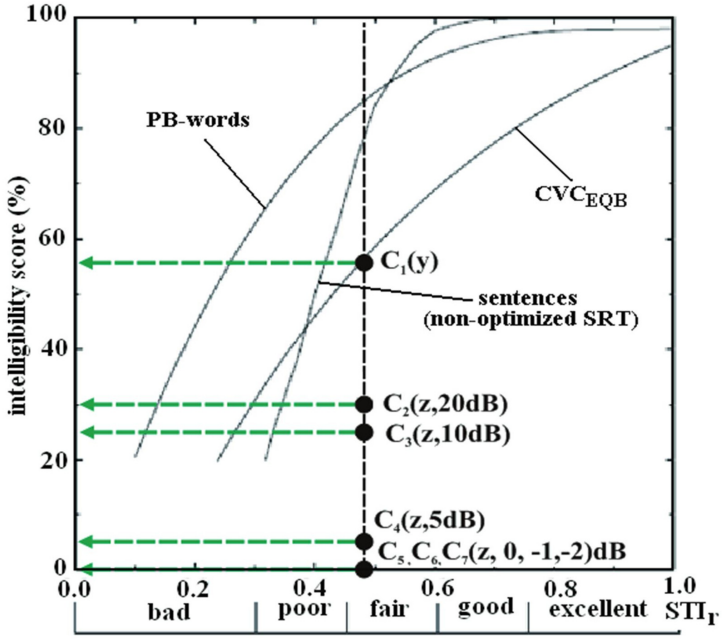

Figure 6. The subjective intelligibility of the sentences by the entire test group in the room with presence of BN8

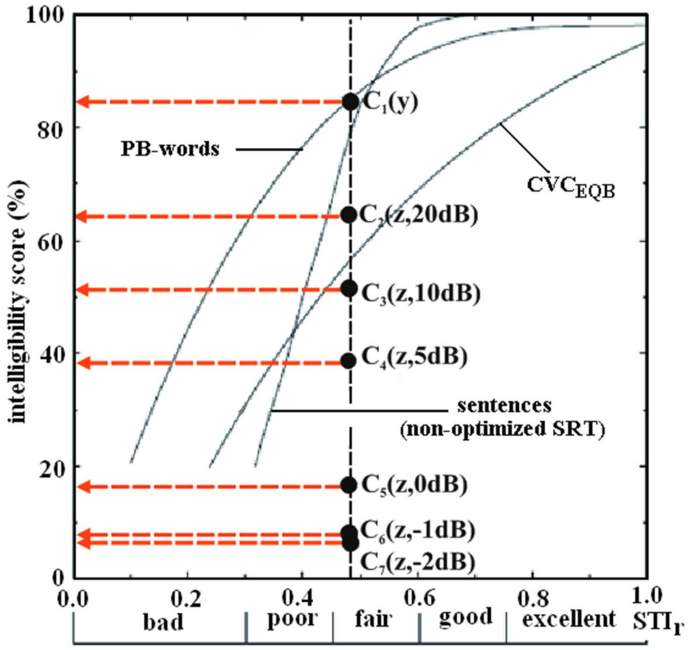

Figure 7. The subjective intelligibility of the PB words by the entire test group in the room with presence of BN8

b) with the influence of **BN8** the male listeners couldn't understand the whole sentences 100% on SNR = {0, -1, -2} dB. The PB words were least understandable $(SIP_{Bw})_{BN}$ = 8 % on SNR = 1 dB.

3) both female and male listeners, with the influence of **BN8**, had the same percentage of understanding: a) for the whole sentences on SNR= 20 dB, $(SI_{WS})_{BN}$ = 30 % (Figure 3.a)) and b) for the PB words on SNR= 5 dB, $(SIP_{Bw})_{BN}$ = 38 % (Figure 3. b)).

4) the best percentage of understanding of both whole sentences and PB words of the whole test group with the influence of BN, is $(SI_{WS})_{BN}$ = 30 %) and $(SI_{PBw})_{BN}$ = 64 %) for SNR = 20 dB (points C2 in the Fig. 6. and 7.). The obtained values with the influence of reverberation are: $(SI_{WS})_r$ = 55 % and $(SI_{PBw})_r$ = 84.4 %, (point C1 in the Figure 6 and 7.) [4]. The estimated values of BN influence on sentences and PB words intelligibility of the whole test group are $(\Delta SI_{WS})_{BN}$ = 25% and $(\Delta SI_{PBw})_{BN}$ = 20.4%. The whole test group showed 0 % of understanding of the whole sentences at SNR = {0, -1, -2} dB (points C_5, C_6, C_7 in the Figure 6.). The least percentage of understanding of the PB words is 6 % at the signal with SNR = -2 dB (point C_7 in the Figure 7.).

4. Conclusion

This paper estimated the degradation level of subjective speech intelligibility with influence of Babble noise and relation to reverberation effect in a room with high reverberation time. For the subjective intelligibility of whole sentences, the influence evaluation value results to be within the range: a) for female listeners ΔSI_{WS} = 22 ÷ 52 %, b) for male listeners ΔSI_{WS} = 28 ÷ 58% and c) for the entire test group ΔSI_{WS} = 25 ÷ 55 %. For the subjective intelligibility of PB words the influence evaluation value results within the range: a) for female listeners ΔSI_{PBw} = 19.6 ÷ 85.6 %, b) for male listeners ΔSI_{PBw} = 21.2 ÷ 71.28 % and c) for the entire test group ΔSI_{PBw} = 20.4 ÷ 76.4 %.

References

[1] Krishamurthy, N., Hansen, J. H. L. (2009). Babble Noise: Modeling, Analysis and Applications, *IEEE Transactions on Audio,Speech and Language Processing*, 17 (7) (September).

[2] Steeneken, H. J. M. (1992). On Measuring and Predicting Speech Intelligibility, *Academisch Proefschrift*, Soesterberg.

[3] Steeneken, H. J. M. (1980). Apsyhical Method for measuring Speech Transmission Quality, *J. Acoust. Soc. Am.*, vol. 19.

[4] Stojanovi, V., Milivojevi, Z., Veli kovi, Z. (2017). Preciznost procene razumljivosti govora na bazi akusti kog impulsnog odziva, *XVI meunarodni nau* no-stru ni simpozijum *INFOTEH*, Jahorina.

[5] Milivojevi, Z., Kosti, D., Veli kovi, Z., Brodi, D. (2016). Serbian Sentence Matrix Test for Speech Intelligibility Measurement in Different Reverberant Conditions, *UNITECH*, Gabrovo.

[6] Stewart, R., Sandler, M. (2010). Database of Omnidirectional and B Format Room Impulse Response, *ICASSP*-88, p. 165- 168.