AURALIZATIONS OF MONOSYLLABIC WORD LISTS FOR HEARING IMPAIRED STUDENTS - A PRELIMINARY STUDY

Konca Şaher[1], Baki Karaböce[2]

[1] Interior Architecture and Environmental Design, Kadir Has University, İstanbul, Turkey;
konca.saher@khas.edu.tr
[2] TUBITAK UME, National Metrology Institute, Gebze-Kocaeli, Turkey
baki.karaboce@tubitak.gov.tr

Abstract
This paper discusses the preliminary results of a research project, which seeks to develop Turkish speech recognition tests by auralizations for hearing impaired students based on monosyllabically structured words. In the context of this study, two sets of 25-items phonetically balanced monosyllabic Turkish words were recorded in the anechoic chamber of TUBITAK National Metrology Institute in Gebze, Turkey. Each monosyllabic word was recorded through a carrier sentence. After the vocal quality, accent and pronunciation in the recordings were approved by qualified audiologists, auralizations of the recorded sentences were developed in an acoustic simulation software (ODEON v12). Listening tests developed from auralizations in three classroom models with varying reverberation times and signal to noise ratios were presented to young adults with normal hearing. The preliminary results of these listening tests are presented and discussed in this paper.

Keywords: Turkish monosyllabic words, auralizations, hearing impaired, classroom acoustics

PACS no. 43.55.Hy, 43.55.Ka

1 Introduction

This paper focuses on preliminary results of a research project titled "Improvement of Classroom Acoustics for the Hearing Impaired" which is supported by The Scientific and Technological Research Council of Turkey (TUBITAK). The project is being carried jointly with TUBITAK National Metrology Institute in Gebze, Turkey.

It has been widely acknowledged by many researchers that the classroom acoustics has an important role on the academic performance of the students. [1-5]. The hearing properties of children are different from that of adults. The children need shorter reverberation time and cannot ignore echoes as well as adults do; children are also more distracted by noise than adults. Children's capability of segregate speech from noise is not that well developed as adults. These have been shown by many researchers. [6-9] Children with hearing impairment are even more vulnerable to acoustic conditions than normal hearing children. There are special schools in Turkey dedicated for hearing impaired children; however, the inclusion of hearing impaired students in the mainstream education is becoming more popular in Turkey. Therefore, it is of interest to investigate acoustic regulations or guidelines for the hearing impaired.

Many countries developed standards for classroom acoustics. The criteria that have the biggest effect on speech intelligibility in a classroom are reported as reverberation time and signal to noise ratio [1-5]. The World Health Organization (WHO) [10] and American National Standards Institute (ANSI) Standard S12.60-2002: "Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools" [11] recommends a reverberation time of 0.6-0.7 sec based on the volume; while Department for Education and Skills: "Building Bulletin 93 Acoustic Design of Schools" (BB93) [12] recommends a reverberation time of 0.8 sec and 35 dB background noise level. It should be noted that these values are described as mid frequency reverberation values. However; many standards do not include values for hearing impaired. The BB93 [12] recommends a reverberation time of 0.4 sec. and a background noise level of 30 dB for the hearing impaired.

One of other organizations that have acoustic recommendations for the hearing impaired are American Speech Hearing Association (ASHA) [13] which recommends a reverberation time of 0.4 sec and this value is described as the average of mid frequency values. ASHA [13] also recommends a signal to noise ratio of 15 dB. However; British Association of the Teachers of the Deaf (BATOD) [14] recommends a reverberation time of 0.4 sec in all frequency range including the low frequency range. The signal to noise ratio recommended by BATOD [14] is 15-20 dB. BATOD [14] recommends a reverberation time of 0.4 sec. in all frequency range; since most hearing impairment occur in high frequencies and therefore hearing impaired might rely on low frequency information to have a better speech intelligibility.

In Turkey, the background noise levels of 30-35 dB are recommended for classrooms in The Turkish Regulations on Assessment and Management of Environmental Noise. [15] However, there are no recommendations for reverberation time or signal to noise ratio in classrooms.

Based on the previous research in the field; this study aims to understand the effect of reverberation time (in all frequency ranges) and the signal to noise ratio on speech intelligibility of normal hearing and hearing impaired students. In order to achieve this, this study also aims to develop Turkish speech recognition tests (similar to the tests made by clinical audiologists) by auralizations for hearing impaired children based on monosyllabically structured words. The subjective listening tests with the monosyllabically words which have phonemes distributed in all frequency range could show the perception of the listeners under different reverberation time and signal to noise ratios.

Therefore, three acoustic models with three recommended reverberation time values from the acoustical standards discussed above were developed based on a real classroom in İstanbul. ODEON v12 room acoustic software was used in the development of all models. The three reverberation time values of the classroom models are 0.8 sec in mid frequency range, 0.4 sec in mid frequency range and 0.4 sec in low and mid frequency ranges. It was indicated in ANSI [13] that normal adults require a signal to noise ratio of 0 dB for high speech intelligibility when listening to familiar speech material. A signal to noise ratio of 15 dB is recommended for hearing impaired by ASHA [13] and BATOD [14]. Therefore, in each model, a signal to noise ratio of 0 dB and signal to noise ratio of 15 dB is to be tested. The auralizations were developed from anechoic recordings of monosyllabically structured words in the anechoic room of TUBITAK National Metrology Institute in Gebze, Turkey

This paper discusses the preliminary results of the listening tests developed from the auralizations of monosyllabically structured words to assess the speech recognition under different reverberation and signal to noise ratio conditions.

2 Methodology

The anechoic recordings were made at the anechoic room of TUBITAK National Metrology Institute in Gebze, Turkey in April 2016. Two sets of 25-items phonetically balanced (PB) monosyllabic Turkish words were recorded. These two sets of 25-items PB monosyllabic Turkish words form one list. This list was taken from the most commonly used PB monosyllabic Turkish words lists, known as PB-300, which was developed by Hacettepe University Audiology Department in 1972. [16] These PB-300 word lists comprise of six lists, each composed of 50 words. Monosyllabic words with two, three and four phonemes were used in each list. The list of two sets of 25-items PB monosyllabic Turkish words which were recorded are given in Table 1 below.

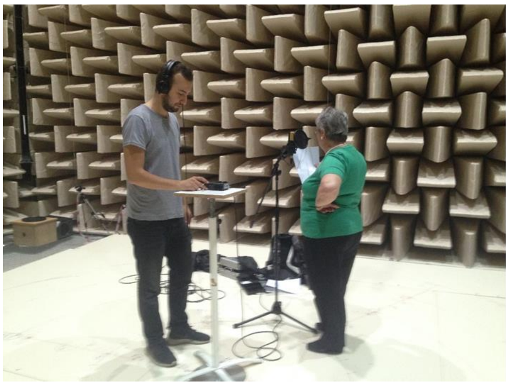
Table 1 – The list of two sets of 25-items phonetically balanced (PB) monosyllabic Turkish words

25-items phonetically balanced (PB) monosyllabic Turkish words lists	
List 1	**Liste 2**
1. KAŞ	1. KAZ
2. TAY	2. TAŞ
3. BİR	3. BEY
4. GÜL	4. GÖK
5. ÇAM	5. CAM
6. BUZ	6. BAŞ
7. TOP	7. TUZ
8. BAK	8. BEZ
9. DİŞ	9. DUR
10. SAĞ	10. SEV
11. KOL	11. KAŞ
12. YAZ	12. YUT
13. ÇOK	13. ÇAY
14. EL	14. AL
15. SAÇ	15. SOL
16. KAN	16. KOR
17. YÜZ	17. YOK
18. GÖR	18. GÜN
19. AT	19. İP
20. DEV	20. DİL
21. YIL	21. YÜN
22. GÖZ	22. DAL
23. UN	23. ÜÇ
24. VER	24. FİL
25. KEL	25. KİR

Each monosyllabic word was recorded through a carrier sentence as the authors of this paper were instructed by expert audiologists. The carrier sentence for each monosyllabic word is as follows:

"Şimdi söyleyeceğimiz kelime kaz".

The meaning of this sentence is: "The word we will say now is goose." The same carrier sentence was repeated for each monosyllabic word. In total 50 carrier sentences were recorded. The recordings were made by one female and one male speaker who has a standard, İstanbul accent in Turkish. Figure 1 below shows the male and female speakers during recordings in the anechoic room and Figure 1 also demonstrates the microphone used during the recordings. There were also simple sentences which comprise of three or four words, which were recorded by the one female and one male speaker. These recordings of simple sentences were intended to be used as background noise in the listening tests. The details and content of these simple sentences are not given here.



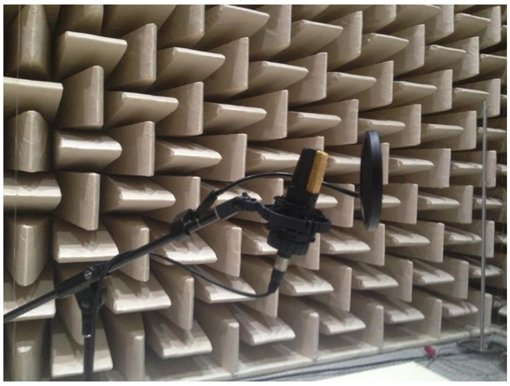


Figure 1 – The speakers in action during the recordings in the anechoic room. (Left figure). The microphone used during the recordings. (Right figure)

After the vocal quality, accent and pronunciation in the recordings were approved by qualified audiologists, three acoustic models with three different reverberation time values (according to the discussed standards) were modelled. The physical classroom model was based on a real classroom in Cibali Secondary School, which is within the Fatih Municipality Area in İstanbul. The dimensions of the classroom and classroom furniture were determined and the materials were examined and noted. In-situ impulse response measurements were not done; however, these measurements are planned to be made to further compare the existing situation with the proposed models.

The classroom was an ordinary rectangle classroom, which has a floor area of 37 m^2 and a height of 3.7 m. The walls and the ceiling were painted, the floor was linoleum and the windows were double glazed. The classroom furniture was also quite ordinary with werzalit (high density wood material) tables and chairs for children and white boards. The interior of the classroom is displayed in Figure 2 below. The classroom was modelled in ODEON as in the existing situation. The wireframe ODEON model of the classroom is demonstrated in Figure 2 below. The following materials from the ODEON material library were assigned to the surfaces:

- Walls: Painted plaster surface;
- Ceiling: Painted plaster surface;
- Floor: Linoleum or vinyl stuck to concrete;
- Door: Solid wooden door;
- The desks (werzalit part): Floating wooden floor;
- The windows: Double glazing 2-3 mm, 10 mm gap.

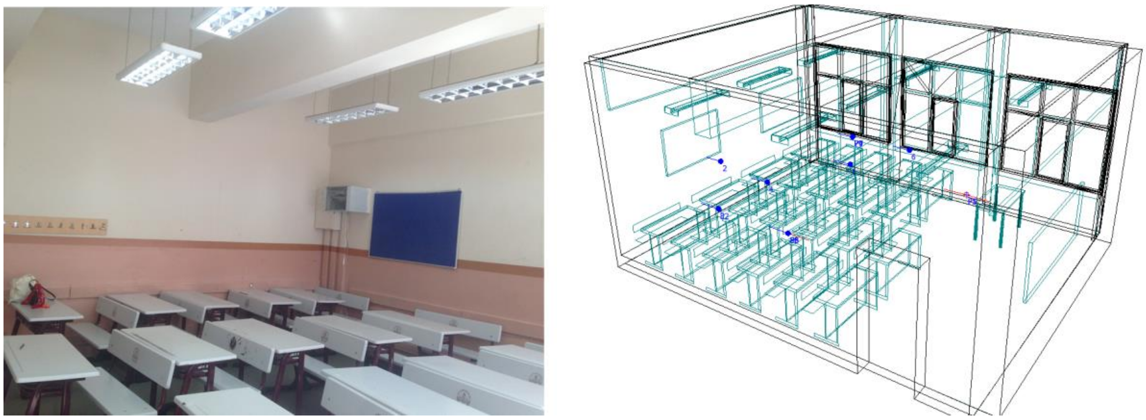


Figure 2 – The photo of the classroom showing the interior materials and the furniture. (Left figure). The ODEON wireframe model of the classroom. (Right figure)

Reverberation time calculations were made for the existing situation in the classroom. A grid response of an omni source was calculated across the floor area at the height of 1.2m. The resulting mid frequency reverberation time was in the range of 2.3 sec. It needs to be noted that the acoustic model was not calibrated with the in-situ measurements, however, the calculated reverberation time was to be expected based on authors' experience and the observations during the site visit.

After the modelling of the existing situation, three versions of the classroom was modelled in ODEON. These classrooms were modelled with three different reverberation time values as shown in Table 2 below:

Table 2 – The reverberation time values to be achieved in each model.

	RT(mid-frequency)	**Based on Standard or Guidelines**
Model 1	0.8 sec. (mid frequency)	BB93 [12]
Model 2	0.4 sec. (mid frequency)	BB93 [12]
Model 3	0.4 sec. (125 Hz -4000Hz)	BATOD [14]

In order to achieve the reverberation time values as per Table 2; new generic absorptive materials were used in the new three models. These absorptive materials introduced were based on Sound Absorber Absorption Class rating which are used to describe the performance of a material as described in international standard ISO 11654 "Sound absorbers for use in buildings: Rating of sound absorption." [17] Absorption classes of materials are categorized from A (good) to E (bad) in this standard.

A *class B* acoustic ceiling was inserted to the whole ceiling area in Model 1 instead of the existing painted surface. The resulting reverberation time in mid frequencies was approximately 0.8 sec. A *class A* acoustic ceiling was inserted to the whole ceiling area and a *class A* acoustic wall material was inserted to the whole back wall in Model 2. The resulting reverberation time in mid frequencies was 0.4 sec. A *class A* acoustic ceiling with very high absorption in the low frequencies was inserted to the whole ceiling area and a *class A* acoustic wall material was inserted to the whole back wall in Model 3. The absorption coefficients of the new absorptive materials in three room models are shown in Table 3 below. It should be noted that the standard ISO 11654 [17] has a lower limit of 250 Hz, therefore, the absorption values in lower frequencies were suggested by the authors and do not necessarily indicate a certain material. Moreover, the absorption values as indicated per material "Class A with low frequency absorption" would be hard to find among general absorptive materials. However, this material was especially used in Model 3 in order to achieve the required 0.4 sec among

125Hz-4000Hz range. The interior view of Model 1 which indicates the absorption characteristic of materials is shown in Figure 3 below. It should be noted that the darker colours in Figure 3 indicates absorbing surfaces.

Table 3 – The frequency response of absorption coefficients of Class B, Class A and Class A with low frequency absorption per model.

Material	125	250	500	1000	2000	4000	8000	Model No
Class B	0.2	0.6	0.8	0.9	0.9	0.8	0.8	Model 1
Class A	0.3	0.7	0.9	0.9	0.9	0.8	0.8	Model 2
Class A with low freq. absorption	0.95	0.95	0.9	0.9	0.9	0.8	0.8	Model 3

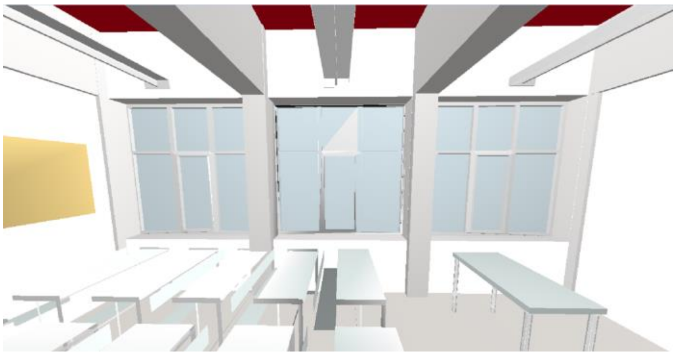


Figure 3 – The interior view of the Model 1 (with reverberation time of 0.8 sec) which indicates the absorption characteristics of materials.

For the listening tests, a scenario which indicates the listener, target source and two additional sources in the classroom was set-up. The location of the each sound source and the listener position is indicated in Figure 4 below. It was assumed in this scenario that the target source was the teacher and two additional sources are to be noise sources (students talking simultaneously).

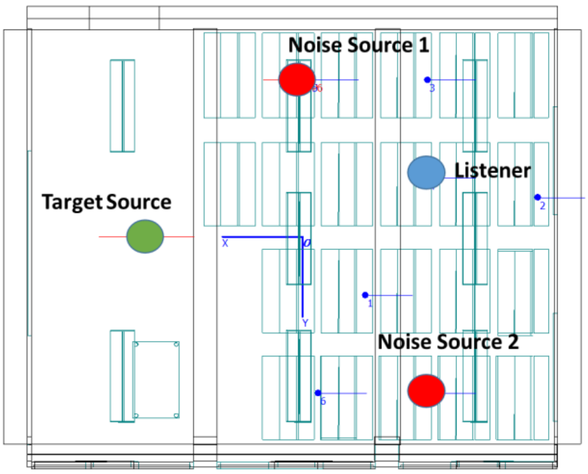


Figure 4 – The location of the target source, the two noise sources and the listener location in the classroom.

Even though 50 carrier sentences were recorded, auralizations of the 25 carrier sentences with monosyllabic words as per Table 1 were made for three acoustic models with two different signal to noise ratio; one is S/N of 0 dB and second is S/N of 15dB. The 25 carrier sentences which were uttered by a male talker was the target source. The anechoic recordings of 25 carrier sentences were convolved with binaural room impulse responses at the listener position. The anechoic recordings of ten simple sentences by a female talker and another ten simple sentences by a male talker were considered as the noise source 1 and noise source 2. (Figure 4)

The anechoic recordings of these simple sentences were convolved with binaural room impulse responses of noise source 1 and noise source 2 at the listener position. These auralizations were to be used as background noise sources when trying to understand the target source. The resulting six different models with three reverberation time values and two signal to noise ratio are summarised as in Table 4 below. These test numbers were to be used for the listening tests.

Table 4 – The description of six different models for six listening tests with three different reverberation time values and two signal to noise ratios.

	RT(mid-frequency)	**Signal to Noise Ratio**	**Test No**
Model 1	0.8 sec. (mid frequency)	0	Test 1 (RT0.8 S/N0)
Model 2	0.4 sec. (mid frequency)	0	Test 2 (RT0.4 S/N0)
Model 3	0,4 sec. (125 Hz -4000Hz)	0	Test 3 (RT0.4L S/N0)
Model 1	0.8 sec. (mid frequency)	15	Test 4 (RT0.8 S/N15)
Model 2	0.4 sec. (mid frequency)	15	Test 5 (RT0.4 S/N15)
Model 3	0,4 sec. (125 Hz -4000Hz)	15	Test 6 (RT0.4L S/N15)

For the each test as per Table 4, 25 auralizations were prepared. Each auralization is a mix of three convolutions at the listener position; first is the convolution of the carrier sentence which include the monosyllabic word (the target speaker); second is the convolution of the simple sentence one (Noise source 1) and; the third is the convolution of the simple sentence two. (Noise source 2) In total, 150 auralizations were made; 25 per each test.

The auralizations in the listening tests were presented through headphones, so special care must be taken to make the levels equal to the levels experienced in real situations. It has been done by adjusting the levels in the computer to the levels the authors believed to be close to the real situation; however; for the next listening tests efforts will be made to calibrate the levels to the real levels.

3 Results and Discussions

The listening tests were made with 15 young adults between the ages of 19-47. Majority of the subjects were between the ages of 22-24; only 3 of the subjects were above 40. None of them reported any hearing problems. Each subject were given two tests; they were given one test with signal to noise ratio of 0dB and one test with signal to noise ratio of 15dB. They were asked to listen to the target speaker carefully and try to understand the monosyllabic word the speaker uttered. The subjects were given a pen and paper where they wrote the monosyllabic word they heard.

Each mono syllabic word written by each subject was checked and a speech recognition percentage was calculated for each test by each subject. The results of these listening tests shown as speech recognition percentage per test is demonstrated in Figure 4 below. The results displayed here are the mean average of subjects for each test. The speech recognition percentage for Test 1 (RT0.8 S/N 0) was 66%; the speech recognition percentage for Test 2 (RT0.4 S/N 0) was 73%; the speech recognition percentage for Test 3 (RT0.4L S/N 0) was 75%; the speech recognition percentage for Test 4 (RT0.8 S/N 15) was 95%; the speech recognition percentage for Test 5 (RT0.4 S/N 15) was 100%, the speech recognition percentage for Test 6 (RT0.8 S/N 0) was 100%.

The results indicate that the speech recognition is increased when the reverberation time is decreased from 0.8 sec to 0.4 sec in the existence of signal to noise ratio of 0 dB. However, it is not very evident that if there is any improvement of speech recognition when the reverberation time of 0.4 sec. is achieved also in the low frequency range. As expected, speech recognition results are almost 100% for all three room models when the signal to noise ratio is 15 dB which is the recommended min value for good speech intelligibility for hearing impaired. The results also indicate a certain trend in the most misunderstood words; however, this is not discussed within the scope of this paper.

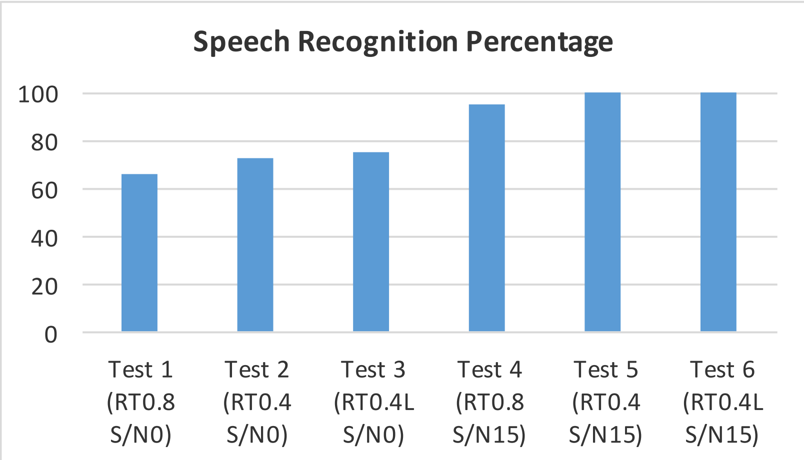


Figure 4 – The percentage of speech recognition among the subjects according to different listening tests.

4 Conclusions

The preliminary results of a research project, which seeks to develop Turkish speech recognition tests by auralizations for hearing impaired students based on monosyllabically structured words are discussed in this paper. In the context of this study, two sets of 25-items phonetically balanced monosyllabic Turkish words were recorded in the anechoic chamber of TUBITAK National Metrology Institute in Gebze, Turkey. Listening tests from auralizations of these phonetically balanced monosyllabic Turkish words were prepared for three acoustic models with three different reverberation time values and for two different signal to noise ratio values.

Listening tests developed from auralizations in three classroom models with varying reverberation times and signal to noise ratios were presented to young adults with normal hearing. The results indicate that the speech recognition is higher when the reverberation time is decreased from 0.8 sec to 0.4 sec in the existence of signal to noise ratio of 0 dB. However, it is not very evident that if there is any improvement of speech recognition when the reverberation time of 0.4 sec. is mainteined in the low frequency range. It is required to increase the number of participants in the future listening tests to be able to have more statistically significant results. Then the extent of the study would be increased to include the listening tests with hearing impaired subjects.

Acknowledgements

The authors wishes to thank Ferda Akdaş for her guidance during the anechoic recordings and assessment of the recorded material; Sezin Nas for the help with the AUTOCAD drawings; Yalın Özgencil for helping with the anechoic recordings; Yahya Burak Tamer for kindly providing some of the recording equipment for the anechoic recordings and TUBITAK colleagues Cafer Kırbaş and Eyüp Bilgiç for their support during the project.

Funding

This work was supported by The Scientific and Technological Research Council of Turkey (TUBITAK) under the 2232 Re-integration Research Grants Program. (The grant number: 21514107-232.01 -230370: Improvement of Classroom Acoustics for the Hearing Impaired.) The project is being carried jointly with TUBITAK National Metrology Institute in Gebze, Turkey.

References

[1] Berglund, B. and Lindvall, T. (1995) Community Noise. Archives of the Center for Sensory Research, 2(1), 1-195.

[2] Institute for Environment and Health (1997) The non-auditory effects of noise. Report R10. http://www.le.ac.uk/ieh/pdf/ExsumR10.pdf

[3] Hetu, R., Truchon-Gagnon, C. and Bilodeau, S.A. (1990) Problems of noise in school settings: a review of literature and the results of an exploratory study, Journal of Speech-Language Pathology and Audiology, 14(3), 31-38.

[4] Evans, G.W. and Lepore, S.J. (1993) Nonauditory effects of noise on children: a critical review. Children's Environments, 10(1), 31-51.

[5] Picard, M. and Bradley, J.S. (2001) Revisiting speech interference in classrooms, Audiology 40, 221-224

[6] Werner, L. and Boike, K.(2001) Infants' sensitivity to broadband noise, Journal of the Acoustical Society of America, 109(5), 2103-2111.

[7] Stelmachowitz, P.G. et al. (2000) The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children, Journal of Speech, Language and Hearing Research, 43, 902-914.

[8] Soli, S.D. and Sullivan, J.A. (1997) Factors affecting children's speech communication in classrooms, Journal of the Acoustical Society of America, 101, S3070.

[9] Johnson, C.E.(2000) Children's phoneme identification in reverberation and noise, Journal of Speech, Language and Hearing Research, 43, 144-157.

[10] World Health Organisation. (1999) Guidelines for Community Noise. http://www.who.int/peh/

[11] American National Standards Institute (2002) Standard S12.60-2002, Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools.

[12] Department for Education and Skills (2003), Building Bulletin 93 Acoustic Design of Schools, www.teachernet.gov/acoustics

[13] American Speech-Language-Hearing Association (1995) Acoustics in educational settings, ASHA Supplement 14.

[14] British Association of Teachers of the Deaf (2001) Classroom acoustics - recommended standards. BATOD Magazine, January 2001.

[15] Çevresel Gürültünün Değerlendirilmesi ve Yönetimi Yönetmeliği no: 2002/49/EC. Çevre ve Orman Bakanliği. (The Turkish Regulations on Assessment and Management of Environmental Noise)

[16] Kılıncarslan A. Türk Dili için Geliştirilmiş Fonetik Dengeli Tek Heceli Kelime Listelerinin Standardizasyonu. Ankara, Hacettepe University, Master Thesis, 1986. (The Standardization of Phonetically Balanced Monosyllabic Word Developed for Turkish Language)

[17] ISO 11654. (1997) Sound absorbers for use in buildings. Rating of sound absorption.