EUROREGIO
BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

Conception of the listening test procedure for quantifying speech intelligibility in Slovak language – a preliminary study

Dominika Húdoková[1,2*], Vojtech Chmelík[2], Lukáš Zelem[2], Daniel Urbán[2], Monika Rychtáriková[1,2]

[1]Department of Architecture - Campus Brussel and Gent, KU Leuven, Belgium.
[2]Department of Materials Engineering and Physics, Faculty of Civil Engineering, STU Bratislava, Slovakia.

* dominika.hudokova@kuleuven.be / dominika.hudokova@stuba.sk

Abstract
This paper describes the methodology used for the conception of the Slovak sentence listening test, suitable for speech reception thresholds (SRT) experiments in quiet and noise. The procedure on choose of suitable sentences in Slovak language is addressed and method on sorting them according to their intelligibility is described. Results shown in this paper are based on a preliminary headphone-based listening test experiments performed on 39 normal hearing female and male human subjects. The task of each respondent was to repeat presented sentence in noise, with individualized signal to noise ratio, tuned for each test person in a pre-test by means of the method of adjustment. The application of the newly developed well-balanced sets of sentences in Slovak language is twofold. Its direct usage can be seen in audiological experiments with severely hearing-impaired individuals and cochlear implant users with Slovak as native language. Secondly, it can serve as a tool for validation of speech related room acoustic quantities, such as *STI* and D_{50} or U_{50} for Slovak language.

Keywords: speech intelligibility, listening test, Slovak language, methodology, sentences

1 Introduction

The speech intelligibility in architectural context is typically performed using single number quantities such as *STI*, C_{50}, D_{50}, U_{50}, etc. [1, 2]. Although the topic of speech intelligibility in architectural context is not a new field of interest, there are several aspects which have not been tackled or fully explored yet e.g. to what extent the mentioned acoustic quantities represent the speech intelligibility of Slovak language? Slovak language, in general, contains large number of consonants and sibilants, which carry lots of ”information”, but carry less sound energy. Thus, we can set a hypothesis that Slovak language might be more difficult to understand in unfavourable conditions when comparing to e.g. Dutch or English. To answer this question, it is necessary to perform a number of listening tests. Several standardized speech intelligibility tests procedures already exist, that has been designed for different languages. Well known are e.g. listening tests based on meaningful sentences using everyday language [3-8], so called matrix test which are specific by means of fixed order of items: name, verb, numeral, adjective, noun etc. [9]. This particular strict content of sentences produces an unpredictable meaning of sentences. [10, 11].
This paper discusses partial results of a dissertation thesis concerning Speech intelligibility in various architectural environments. The main aim is a detailed discussion on methodology for the design of the sentence listening test for Slovak language.

2 Methods

2.1 Criteria for selection of sentence database

Altogether, 800 ordinary Slovak sentences were suggested. The choice of the sentences was based on the following presumptions: (a) to be short enough to repeat. It means that sentences contain minimum 2 and maximum 12 words. (b) Sentences represent conversational speech. There are no scientific phrases, difficult names or specific brand names. (c) None exclamations or questions. (d) Sentences should not to be too unessential. (e) There are no proverbs, to avoid the conceit. The criteria for designing Slovak sentence test material were inspired by van Wieringen and Wouters [7].

2.2 Psychoacoustic method used for listening test

Psychophysics generally focuses on detecting the relationship between a physical stimulus and its perception and on detecting the phenomenon [12]. In psychoacoustic listening tests, we rely on people's answers. Therefore, statistical analysis is essential in data analysis. Statistically relevant sample of people is one of the starting conditions. The first method used in our experiment is called method of adjustment, see in the Figure 1 below. In this method, subject is asked to control the level of the stimulus and keep it changing until the stimulus is narrowly detectable against the background noise [12].

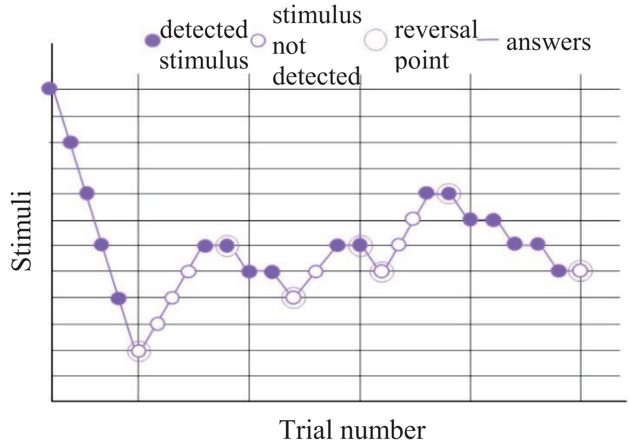

Figure 1: Graph illustration of adjustment method.

3 Listening experiments

3.1 Test subjects

In total, 39 human subjects (20 male and 19 female) participated in the listening test experiment. Their age was between 25 and 61 years. The average age is 33 years, the modus value is 27 and the median value is 28 years. Every individual has declared that he/she did not suffer from any hearing impairment and all of them had normal hearing (<20 dB HL hearing loss range for octave frequencies from 125 Hz to 8000 Hz, measured in the laboratory). Each test subject, that participated in the experiment, received the same explanation from the test operator. They were also explained that the psychoacoustic experiments are not an exam with correct or incorrect answers or a competition on performance and that the most important is to give honest and true answers.

3.2 Stimuli

The recordings of the 800 chosen sentences were performed in the Acoustic lab – the semi-anechoic room at STU Bratislava, Faculty of civil Engineering, Department of Materials Engineering and Physics. They were spoken by female person and recorded at normal speech level, regular speed and spoken in monotonic way without emphasis on any of the words. During the recordings, the speaking person stood in front of microphone in the empty semi-anechoic room alone. The recorded signals were sampled at 44100 Hz (16 bit A/D converter). In editing process, 30 ms cosine window was applied at the beginning and the end of each sentence to avoid click sound.
The whole sentence material was stored in a ".wav" format on disk of the computer. The masking noise necessary for performance of the tests was based on the long-term average speech spectrum of randomly picked 40 sentences (the duration of sample was 2 minutes) of 800 sentence the database. It was a stationary noise with a spectrum of the female voice (used for recordings of the signal in this experiment) thus, a so called spectrally matched masker.

3.3 Set-up

The set up consisted of the computer with home-made, Visual Basic based listening test application, the sound card (Focusrite Scarlett 6i6), two pairs of headphones (one used by test subject and one by operator) and a microphone. The test subject was seated in the semi-anechoic room in the Acoustic lab (Figure 2) and performed listening tests by means of calibrated open headphones (Sennheiser HD650) with a flat frequency response. The headphones were calibrated using analyser Norsonic 140. The sentence material was played from a computer into the headphones via sound card. The test subjects communicated with operator via dynamic microphone, e.g. when repeating the sentences during the listening test, that they have heart via headphones. The test operator was seated in the control room and listened to the responses of people via the second headphones (Figure 2).

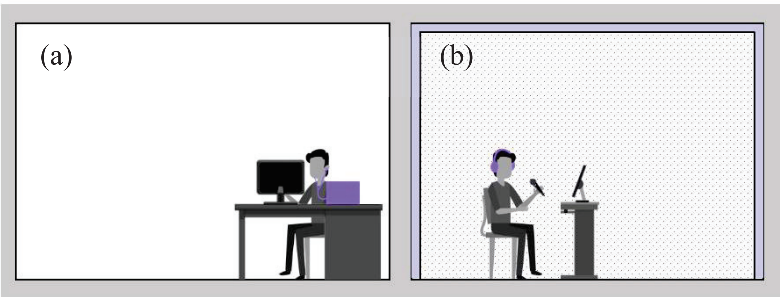

Figure 2. The schematic sketch of test rooms ((a) the control room, (b) the semi-anechoic room).

Test No.1
The first test served for estimation of the approximate SRT (Speech Reception Threshold) necessary for setting up the correct S/N ratio for the Test No.2. The psychoacoustic method used in the test No.1 was adjustment method and consisted of 40 sentences. Each one of the 40 sentences started at S/N ratio = -20 dB. The task of the test subject was to reduce the noise in 1 dB steps until he/she was certain about the meaning of that particular sentence, what means he/she could understand the whole sentence. Each sentence was played in the loop so that the test subject had the opportunity to listen to it as many times as needed. When it was done, the test subject could continue to another sentence. When the Test No.1 was finished, the signal to noise ratio was calculated from the average of all (40) tested sentences.

Test No.2 (main test)
The second test consisted of 760 sentences. These sentences were masked by the noise signal with fixed S/N ratios obtained in the Test No.1, and thus individual for each test subject. The task of a test subject was to listen each sentence in the fixed S/N ratio and to repeat it. Each sentence was played once. The operator seated

in the control room was present all the time and has evaluated answer, i.e. if the sentence was repeated correctly (i.e. each word of the sentence was correctly repeated) and noted the score. The operator could hear each sentence (played to test person) without presence of masking noise. For double check, the sentence which was played to a test person was at the same time also displayed on the screen of the operator. To avoid fatigue and errors in the test, participants were allowed to ask for a break anytime.

4 Results and discussion

The results of the Test No.1 are shown in the Figure 3. They show that the S/N ratio of the most (17) participants was -6 dB. Only two participants have adjusted the S/N ratio to -8 dB.

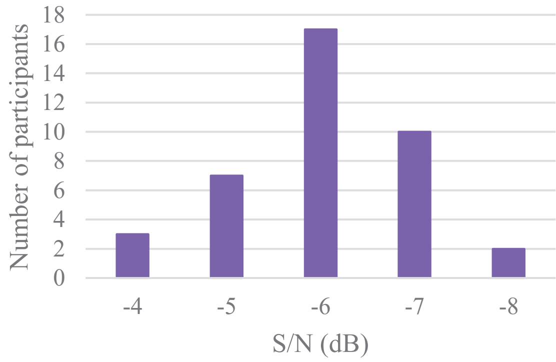

Figure 3: The histogram of S/N (dB).

The results of the Test No.2 were gathered in the matrix, where correct answers were indicated as “1” and incorrect as “0”. Doing so, each sentence got a score, which was later expressed in percentage of correct answers. Sentences were later sorted according to the easiest to most difficult in a form of a histogram. The global result can be seen in the Figure 4.

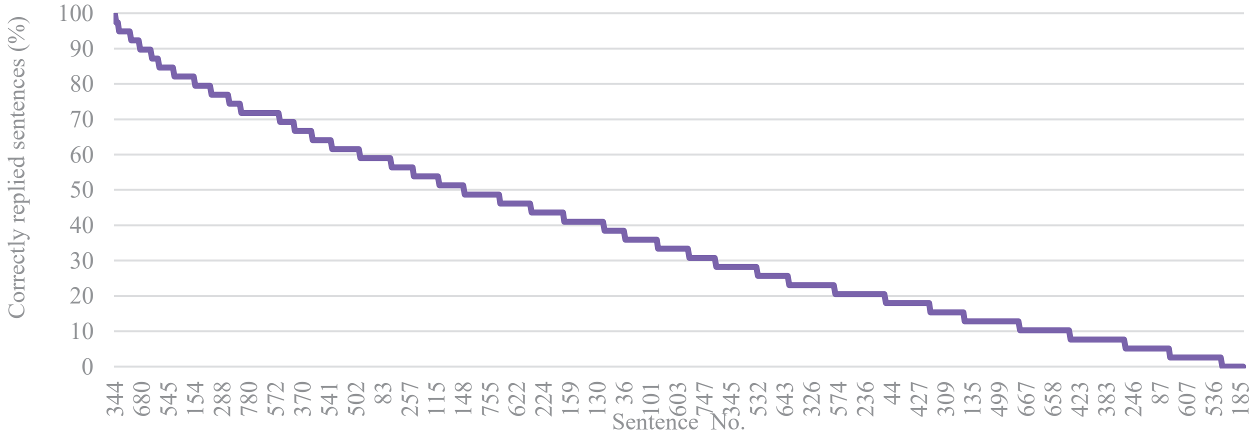

Figure 4: The percentage of the correctly replied 760 sentences. Showing the course of most simple to most difficult sentences.

The percentage of success rate includes whole range from 0% to 100%. We can see by looking at the Figure 4 that the whole sentence material included higher amount of too difficult sentences in comparison with too easy sentences to repeat. This is represented by slowly decreasing slope of presented histogram.

In order to determine the set of equally difficult sentences to understand, the following method was used: (1) the standard deviation of the percentage of correct replied sentences was calculated (STD=26%). (2) the centre of the whole percentage range was indicated (50%). (3) the standard deviation was added to centre of percentage range as well as subtracted from the centre to obtain both the top and the low border of the range (50%+26%=76% it is the top border; 50%-26%=24% it is the low border). We can say that the equally difficult sentences are between 24% and 76% what means 52% of the sentence material.

If we look into the results deeper we can conclude the following: The levels of signal to noise ratio oscillated between -4 to -8 dB, but the level -6 dB was detected in the most of tested participants (Figure 3). In comparison with van Wieringen and Wouters (S/N = -8 dB; -10 dB) [7] we obtained lower average value of S/N. This can have several reasons, among which one can be caused by the fact that the Slovak language due to its large consonant content, is more difficult to be understand under presence of noise. However, this would need to be confirmed by a dedicated study. The majority of participants were adults between 25 and 35 years. Based on our observation, we can claim that age of participants had not affected the results of signal to noise ratio. (Figure 5). Moreover, the age has not affected the number of correctly repeated sentences either.

The range of correctly repeated sentences over participants was between 71 and 472 (Figure 6).

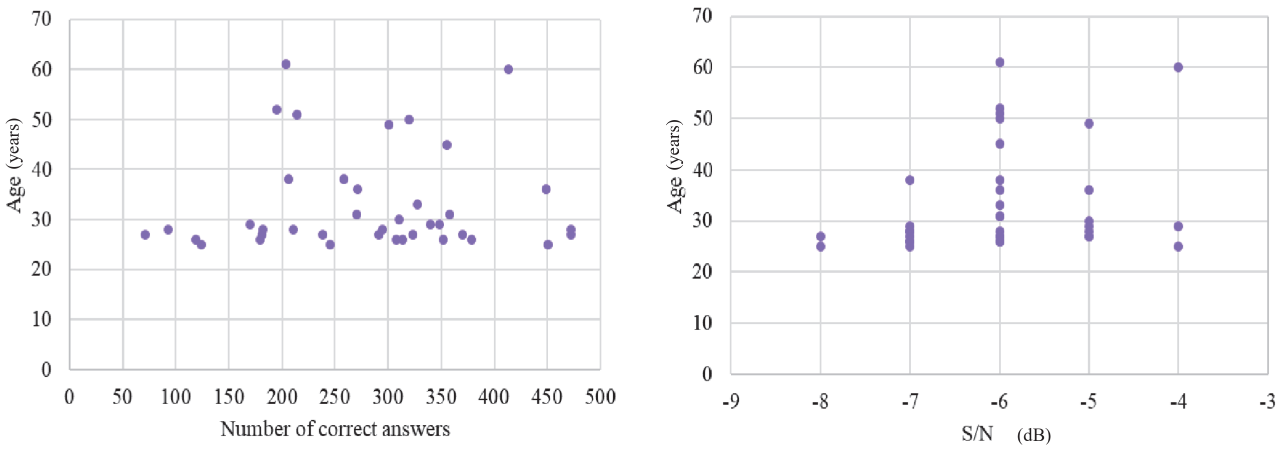

Figure 5: Correctly repeated sentence and S/N ratio in a function of the age (of participants).

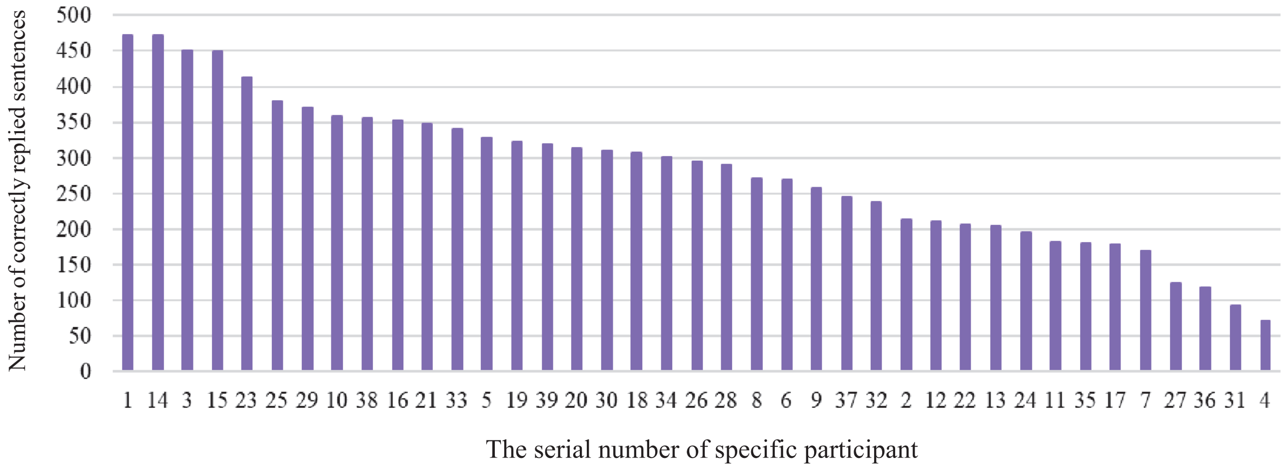

Figure 6: The quantity graph of correctly replied sentences of each test participant.

5 Conclusions

In this paper a sentence material consisting of 800 Slovak sentences was created following the rules based on the work of van Weiringen and Wouters [7]. According to the results based on laboratory listening tests, with 39 test subjects, the ranking of proposed sentences from the easiest to the most difficult has been made. A deeper look into the sentences and explanation upon their difficulty is given in [14]. Altogether 376 sentences have been indicated as equally difficult.

Acknowledgements

The authors would like to express a gratitude to all participants of listening tests.

References

[1] ISO 3382-1 Acoustics – Measurements of room acoustic parameters, Part 1: Performance Spaces.

[2] IEC 60268 Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index.

[3] Plomp, R. – Mimpen, A. M.: Improving the reliability of testing the speech reception threshold for sentences, *Audiology*. Volume 18, Issue 1, p. 43-52, 1979.

[4] Nilsson, M. – Soli, S. D. – Sullivan, J. A.: Development of the hearing in noise test for the measurement of speech reception thresholds in quite and in noise, *The Journal of the Acoustical Society of America*. Volume 95, Issue 2, 1994.

[5] Versfeld, N. J. – Daalder, L. – Festen, J. M. – Houtgast, T.: Method for the selection of sentence materials for efficient measurement od the speech reception threshold. *The Journal of The Acoustical Society of America*. Volume 107, February 2000.

[6] Wong, L. L. N. – Soli, S. D.: Development of the Cantonese Hearing in noise test (CHINT), *Ear and Hearing*. Volume 26, Issue 3, p. 276-289, 2005.

[7] van Wieringen, A. – Wouters, J.: LIST and LINT: Sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands. *International Journal of Audiology*. Volume 47, Issue 6, p. 348 - 355, 2008.

[8] Ozimek, E. – Kutzner, D. – Sek, A. – Wicher, A.: Polish sentence tests for measuring the intelligibility of speech in interfering noise, *Itnernational Journal of Audiology*, Volume 48, Issue 7, p. 433-443, 2009.

[9] Kollmeier, B. – Warzybok, A. – Hochmuth, S. – Zokoll, M. A. – Uslar, V. – Brand, T. – Wagener, K. C.: The multilingual matrix test: Principles, applications, and comparison across languages: A review, *International Journal of Audiology*. Volume 54, Issue 2, 2015.

[10] Hagerman, B.: Sentences for testing speech intelligibility in noise, *Scandinavian Audiology*. Volume 11, Issue 2, p. 79-87, 1982.

[11] Panocová, R. – Gregorová, R.: Designing the Slovak matrix sentence test, *International journal of applied languages and cultural studie*. Volume 2, Issue 2, 2019.

[12] Rychtáriková, M.: Psychoakustické testy v stavebnej akustike. 2015 ISBN: 9788097199364

[13] Gescheider, G.: The Classical Psychophysical Methods. Psychophysics: the fundamentals (3rd ed.). Lawrence Erlbaum Associates. ISBN 978-0-8058-2281-6, 1997.

[14] Hudcovičová, B. – Petrášová, M.: Linguo-Acoustic acpects of speech, *EUROREGIO/BNAM2022*. Aalborg, Denmark, 2022.