Single Number Speech Intelligibility Measurement Methods to Support the Objectification of Educational Research

R. Drtina, J. Sedivy

Abstract**—**Article takes a look at single-digit measurement methods of speech intelligibility as objectification promote educational research that is concerned with the quality and effectiveness of education process. The results to be as objective as possible, it is necessary to determine the initial terms of quality of information transfer. For evaluating the quality of speech transmission there is a one-digit number of measurement methods. The above-outputs but are not directly comparable or coordinates results in relatively wide zones. Clarity of speech should be a key requirements of the educational process. Still remains the most accurate method logatomic tests - files statistically selected syllables) that semantically related and can thus infer from the context of the sentence as a normal continuous speech**.**

Keywords**—** clarity, communication, information, quality transfer, speech transmission index, transfer

I. Introduction

DESPITE the escalating requirements for multimedia support teaching, improving the quality and effectiveness of education is the formation of the working environment in the context of the requirements of ISO 9001, Art. 6.3 - Infrastructure and 6.4 - Working environment and decrees MMR 137/1998 Sb. and MZ No.108/2001 Coll. is essentially zero. In the case of educational research and are never objectively describe the input conditions. These can significantly affect the results (e.g. evaluation of listening to language teaching). In the last years we randomly example.

Good working environment educational institutions, schools, workplaces and domestic (e.g. In the use of e-learning courses) is subject to a number of hygiene and ergonomic requirements, it is necessary to ensure coherent optical and acoustic performance while achieving maximum possible transmission quality optical and acoustical information and minimize visual and auditory fatigue. In terms of research carried out must be ensured and necessary repeatability in terms of ISO 5725-1 [8] ISO 5725-2 [9] ISO 21748 [10] ISO 2602 [11] ISO 2854 [12]

R. Drtina, University of Hradec Kralove, Faculty of Science, Department of Technical Subjects, Rokitanskeho 62, 500 03 Hradec Kralove, Czech Republic (phone: +420 493331171; e-mail: rene.drtina@uhk.cz).

J. Sedivy, University of Hradec Kralove, Faculty of Science, Department of Informatics, Rokitanskeho 62, 500 03 Hradec Kralove, Czech Republic (phone: +420 493331171; e-mail: josef.sedivy@uhk.cz).

II. Methodology and Information Technology

Likewise, vision and hearing (video and audio) receptors are still the dominant sense in mediating knowledge of the surrounding world. Although at this year's Consumer and Computing International Consumer Electronics Show head of the US subsidiary Bosch, Werner Struth, said that the "analog world as we know it ends" [1], our perception still remains analog as well as continue to persist transfer information analogue transmission channels which may interfere with, among others, restrictions resulting from digital technology [2].

Various educational research, studies on the quality and efficiency of the educational process should do so as objectively as possible and in connection with the implementation of quality management to take into account the requirements of ISO 9001, which in Art. 6.4 - Working environments states: "The organization shall determine and manage the work environment needed to achieve compliance with the requirements for products. The term work environment relates to the conditions under which the work is done. These conditions include physical conditions, under-conditions and other environmental factors (such as noise, temperature, humidity, lighting or weather). "[3]

Infrastructure states that "The organization shall determine, provide and maintain the infrastructure needed to achieve conformity to product requirements" [4]. Under the circumstances, then infrastructure includes:

- buildings, workspaces and related technical equipment,
- process equipment (both hardware and software), and
- support services (e.g. Communication or information systems).

If we identify, from a technical point of view, for acoustic information transmission didactic principles of clarity in the teaching process speech intelligibility, then we can, while respecting acceptable losing clarity of consonants at ZSS 5% derive requirements for acoustic transmission channel [5].

The primary prerequisite acoustic communication is very good logatomic cognizable and its derivative sentence intelligibility. Given that in professional courses is acoustic information transfer based on communication pupils primarily

unfamiliar technical terms, we can rightly claim classrooms logatomic cognizable better than 97%. This can be easily verified using logatomic tests that are processed in tables for each language. Its verification and subsequent statistical evaluation is feasible without major problems means each school. Based on the results can then decide on the need and extent necessary acoustic treatment of the classroom and its technical equipment [6].

Consistent application of the golden rule didactics is for obvious reasons a source of higher demands on teachers. In practice this means that it is important not only to prepare acoustically provided information but also to adapt its communication process pupils with regard to their maturity, but also on the surrounding environment [7]. The speed of speech must adapt perceptual abilities of students (normally recommended speech rate is approximately five syllables per second), to ensure proper articulation and speaking loud enough for all students to have a level playing field perception. Doing their own (ambient) noise classrooms, reverb and other disturbances significantly impair intelligibility. In this context, it is often underestimated disturbing noise in computer rooms [8].

Workers electrical engineering laboratories of the Department of Technical Subjects of Faculty of Education, University of Hradec Králové many years solves the problem of objectification measuring intelligibility in classrooms in collaboration with colleagues from acoustic laboratories FEL West Bohemia in Pilsen. The results of recent researches show often a stark difference between the Speech Transmission Index, measured by a standardized method of bite, insensitive to sources directivity and frequency characteristics of the transmission channel and the results obtained by logatomic tests. The aim of the research projects carried out in recent years, is to verify the measurement methods so as to ensure the correlation of the signal measurement method Stipe results of valid tests logatomic so ideally correlation coefficient ρCIS, LOG approached one, and that the measured values take into account the directionality of sources a bandwidth of.

III. Repeatability Tests and Security Conditions Research

In order for pedagogical research in the field of monitoring the effectiveness of training and evaluation of the quality and success of distance courses, in which an audio material, able to clearly define the input acoustic conditions of the working environment (this applies to both the spatial acoustics of lecture halls and classrooms, and home environment acoustics in the use of distance forms of teaching), and to be assured of technical reproducibility of the experiment in terms of standards [8-12], it is necessary to have a reference measurement signals.

Problems with the inconsistency of the results of different measurement procedures and methods are nothing new. Already researches conducted in the 50s and 60s of last century, a number of research organizations have shown different results of speech recognizability tests depending on the language. For example [9] reported that the differences between Czech and English, ceteris paribus, reaching up to 15% in favor of Czech, but not the Slavic languages are the results compared marketable. In the frequency range 600-2000 Hz is the difference between the recognizability of syllables Czech and Russian to 22%, again in favor of the Czech.

Both authors were well aware of the complexity of the issues and evaluate the clarity of their statements in that document that "if we emphasize the need for objective measurement methods in linguistics and especially in experimental phonetics, we also note that the very use of modern measuring equipment is unable to solve definitively any linguistic problem. We are fully aware that devices are only auxiliary equipment, the results of which must be individually processed and controlled by subjective methods. Although we expect in future not only with mechanization, but also automating highly complex measuring tasks and their evaluation, the final factor that will always remain one. It is due to the fact that the linguistic issues in terms of objective reality, which is always perceived subjectively by man "[10] is valid even after more than fifty years, in the era of digital technology and sophisticated measuring systems.

Also, previous research results evaluation method clarity, articulation index (speech intelligibility index - SII, ANSI S3.5 - 1997), which is used by university departments e.g. To evaluate the clarity in vehicles [11], which is implemented in the system of Czech technical standards, show that the method SII has considerable limitations, especially in cases where it is most disturbing noise below the reference areas of speech, articulation index is approaching 100% and above-the results are comparable with the results of standard listening tests.

Determining the articulation index is not according to the standard method suitable for areas with a low level of self-noise and thus provided the results can not be considered fair. Therefore, introducing the concepts of "open articulation index" and "closed (Traditional) articulation index". Contributions to open speech articular index are not limited to the lower limit of the speech band. It is shifted to 0 dB and articulation index reaches values over 100%. Theoretically, the maximum attainable then approaching 225%.

IV. Logatomic Tests

The analysis of the results of measurements and signal spectra also demonstrated insensitivity standardized methods for frequency and directional characteristics of the emitters, which was at the beginning of the study rather negative findings highlighted the need for us to return to practice proven, time-consuming, but very sensitive method of detection achievable clarity - logatomic tests. We have unsuccessfully searched for the original recordings, which arose in the 60s and 70s for VÚZORT, AZD or Language Institute, Czech Academy. Most likely they are unique records irretrievably lost.

In cooperation with Czech Radio Hradec Králové (Radio

HK) Therefore we have produced a new version logatomic tests for the detection of clarity (or more precisely - the recognizability of the syllable) in classrooms and auditoriums. At this point, we consider it necessary to emphasize that the reference tests must be perfect from the technical and phonetic [11]. It is therefore not possible to create in amateur conditions. The record shall be drawn up without corrective actions, without adjustment dynamics and must meet the requirement so. Voice format. This means that the signal from the reference monitor (in our case the reference emitter NTi Talk Box) must be the same as if at a given location within a particular speaker spoke.

Before the start of production, we measured the basic parameters of the study and transmission paths. Radio HK has anechoic space called "plein air". Therefore, in the studio built full damping panels (Fig.1). The average reverberation time of study is in the range of 270 ms of speech (Figure 2) and the radius of reverberation for omnidirectional radiation is 103 cm. Pro speaker is then the reverberation radius 162 cm distance and reverberation than 8 dB. Reverberation radius is so beyond the exterior wall panels and damping due to high distance is the dominant direct signal.

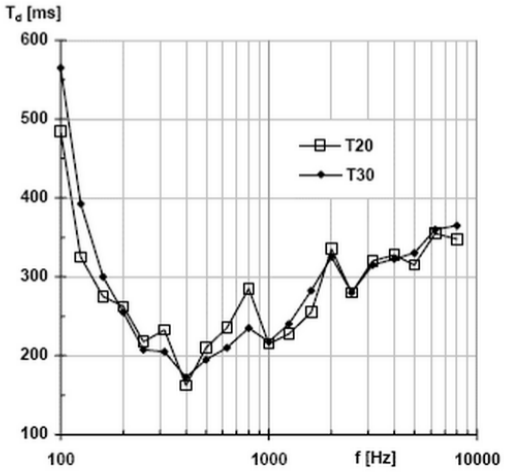

Fig. 1 The reverberation time of studios

With reference emitters NTi Talkbox test was conducted around the transmission channel. At the workplace moderators (Figure 2) was placed reference source and an alert has been measuring signals bite, white noise and pulses.

Neumann U87 microphone sensing is a professional studio standard. The recording was not called. Pop-filter with an omnidirectional polar pattern. The transfer characteristic is in the range of 100 Hz to 10 kHz (range NTi Talkbox) equals, speech transmission index Stipe gives the maximum possible value STI = 1.00 and A + quality, which is excellent speech, which can be in studio settings rightly expect.

For applications in educational research, we chose a compromise between the length logatomic test and achievable precision. Hundred syllables test has a resolution of 1%, and is divided into five groups of twenty syllables. When the cadence of 0,5/s i.e. One syllable in two seconds, with a break of 10 s between groups, one test takes less than 4 minutes. To listeners in repeated tests could remember the order of syllables are in various forms of tests swapped to the entire group, so each syllable.

The production tests involved a team of five employees of Czech Radio HK. Moderators Jakub Schmidt, Lada Klokočníková, Karel Sladký, sound engineer and director Leos Sedlacek Pavel Krejci. As we learned during manufacturing, currently director Tailor filmed in the seventies logatomic tests for AZD.

V. Variability of Logatomic Tests

Collaboration with professionals, particularly in the production of a key test of audio material, consider not only commonplace, but a necessary condition for objectification of input values pedagogical research whose content is aural component of the learning material. To get near the practice tests were filmed in three versions for voice, lower male voice, higher male voice and a female voice. Professional presenters under the leadership of literary director for us were the guarantee of a precise pronunciation of speech therapy without defects [12].

Equally important is a stable rate of speech, a constant cadence of logatoms and also neutral vocal performance without modulation or intonation changes. The structure and quality of the recorded material allows seamless combination of any 20 syllables groups of up to 120 tests for one vote. The combination 20 syllables groups from all three votes, then theoretically allows you to create over 360 thousand tests.

In addition to the basic transmission Kana testing and monitoring the impact of corrections, directivity emitters, transmission speed (bitrate) and associated digital signal degradation, we comparative method to determine the effect of imperfect articulation, speech defects, or improper technical processing of sound recording on the achievable intelligibility under otherwise identical conditions, transmission . Just to author learning material loaded the same logatomic test and comparing the results we get peace deterioration due to vocal intelligibility of speech, recording conditions, combination of both factors.

VI. Educational Experiment

In the framework of research project specific research we verified experimentally the influence of vocal expression and listening conditions. The task was to make students logatomic test in school environment cabinet or home office. To record the pillar was used electret microphone, commonly used for communication via Skype and a computer with standard configuration Win7. Very soon students know that record just one 20 syllables stretch in one sequence is not so simple.

Listening was the monitors RS-711H in the near field (60 cm, STI = 0.99) and built-in speakers Sony Vaio laptop (60 cm, STI = 0.93). The results are shown in Table 1.

Table 1 The comparative test results

	record production	radio	home
listening device - the source format		intelligibility	
monitor RS-711H	PCM 16 bit/44,1 kHz	100 %	89 %
Sony Vaio	PCM 16 bit/44,1 kHz	100 %	87 %
monitor RS-711H	MP3 56 kb/s	100 %	88 %
Sony Vaio	MP3 56 kb/s	98 %	87 %
monitor RS-711H	PCM 8 bit/16 kHz	94 %	87 %
Sony Vaio	PCM 8 bit/16 kHz	91 %	85 %
monitor RS-711H	PCM 8 bit/8 kHz	81 %	74 %
Sony Vaio	PCM 8 bit/8 kHz	78 %	73 %

The results in Table 1 show that the differences for high audio quality may reach tens of percent. It is quite logical that bad transmission and listening conditions contrary, differences partly blurred. The values shown are for illustration purposes only and benchmark results can not be generalized small data file. We just want to show that the production of audio materials for teaching, we should pay due attention. The fact remains that unlike logatomic the recognizability of the sentence intelligibility in adverse conditions is significantly higher because of the lack of conjecture out of context sentences.

VII. Spectral Composition of measuring Signals

Because our main goal is to find the relevant relationship between measurement methods and results STIPA logatomic tests, respecting the width of bandwidth and radiation characteristics, we performed spectral analysis files logatomic tests and signal STIPA.

The following graphs show the peak values of each spectral measurement signals. These are different from the conventionally reported power spectral density. The energy potential of formant frequency band, and sibilance is relatively small. Their immediate acoustic performance but significantly higher and ensures speech intelligibility.

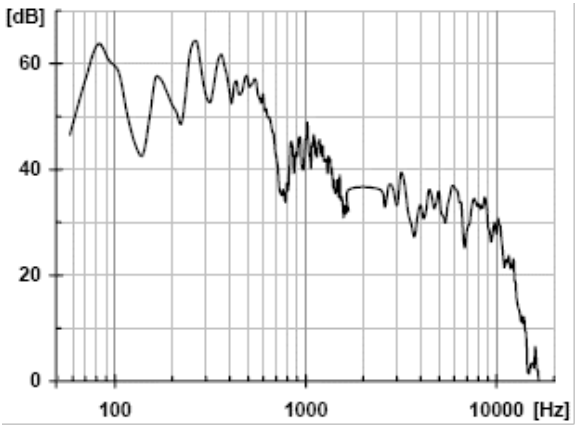

Fig. 2 Spectrum of lower male voice (K)

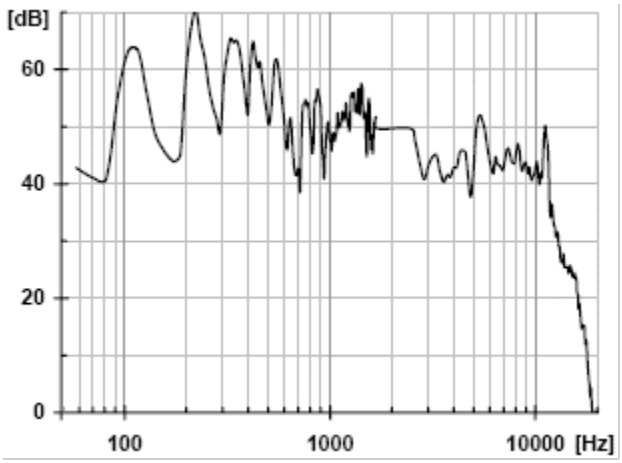

Fig. 3 Spectrum of higher male voice (J)

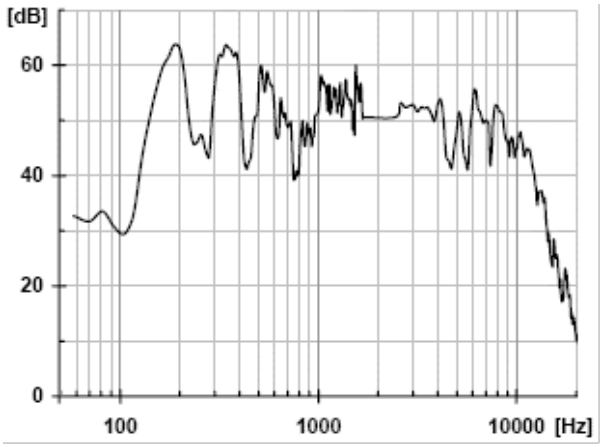

Fig. 4 Spectrum of female voice (L)

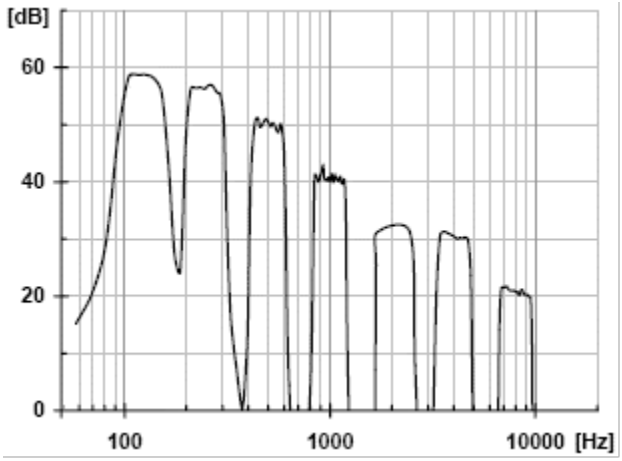

Fig. 5 Spectrum of STIPA signal (S)

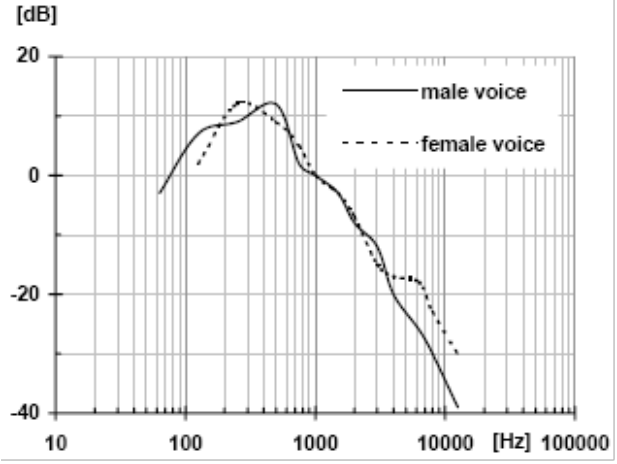

Fig. 6 S Power spectrum of speech

A comparison of the measured spectra can be traced to the female and the male voice has a higher instantaneous power balanced between basic tone, formant areas and sibilance. In contrast, the lower male voice has a stronger fundamental tone and almost perfectly agrees with the spectrum of the measurement signal chips, which is validated only for male speech.

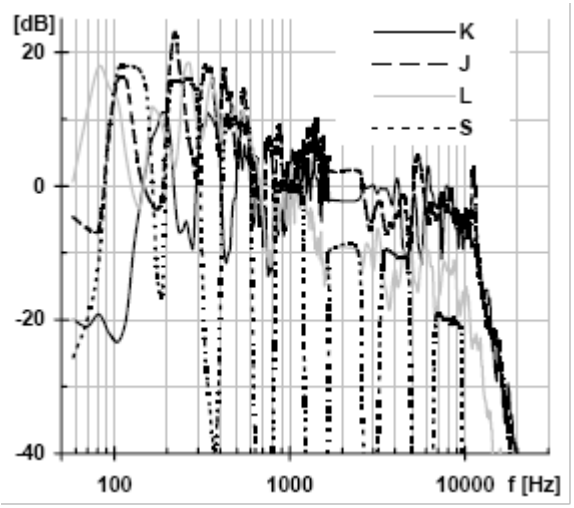

Fig. 7 Compare of the spectra of measuring signals

VIII. Another Direction of the Correlation Research

Generally, the objective correlative research to determine the closeness of relationships, relationships and dependencies between variables. The aim of our research is to find a relationship between the width of the transmitted bandwidth directional characteristics of radiators, measurements by Stipe and actually found logatomic recognizability.

Conducted research on Czechoslovak (1990) and Czech research, which focused on setting intelligibility using logatomic tests do not give cause for optimism. According to us, the collected data is the most extensive research dating to the 50th-70th 20th century. On the direct question of the Czech Language Institute, we were told that after the abolition of the department of phonetics and migration institute today unfindable not only the original audio recordings, as well as research reports or documents the once-planned publications. And there were several research projects with a considerable range of processed materials. Current phonetics monitors, among other things, especially the modern trends of speech recognition and synthesis for the needs of the development of speech technologies. Evaluation of transmission characteristics of the reproductive systems are diverted from qualitative testing and quantification favors the description of one of the measuring methods, taken from ANSI.

Education is to some extent a specific area in which pupils and students widely presented primarily unfamiliar concepts. Therefore, we believe, particularly in view of the fact that schools are still accepting 50-80% of the information hearing that perfect clarity, not just sentence or a word, but just syllable, is a prerequisite for good distinguishable phonemes. This minimizes erroneous interpretation of subjective information received acoustic transmission channel and create favorable conditions for predictive function of interpreting speech as the primary bearer of information (for example, when creating and defining new concepts and relationships).

Since the beginning of the research we are facing a problem that is very difficult to build on current research or research from recent past. We have to go back to good works classics Czechoslovak acoustics and the complement of related contemporary knowledge of phonetics. A necessary fact is the fact that the acoustics of classrooms has always stood and still stands outside the main concern acousticians and engineers. With the development of distance education, e-learning and trying some workplaces to distribute recordings of lectures (often with questionable technical quality), we should seriously address the issue of domestic work environment and the quality of the transmission channel in which the audio and video learning materials spread.

IX. Signal Speech Samples - Corpus Monologue

Body monologue contains audio recordings of the sound archives of Czech Radio and is primarily designed primarily as a source of research pronunciation in Czech media. Among other things it serves but also to those who are engaged in analyzing sound, i.e. the phonetic, Czech plane.

Our interest is focused on the quality of sound recordings. In the monologue corpus we found a recording of a female voice, which resembled our logatomic tests. We evaluated approximately three minutes of continuous text and come to the unpleasant discovery. Stored record is not a record of pure voice, but broadcasting signal processed by the department.

In the following chart, the peak value of the frequency spectrum of the recording of the corpus monologue. Of course the level is visible lift base tone (up to +30 dB in the band 200-600 Hz), faint sibilance in the range of 5 kHz and a bandwidth limiting over 12.5 kHz stereo signal modulation. Level difference among neighboring frequencies are relatively

small and as a result, the frequency characteristic of the voice is relatively smooth. It is quite obvious that the resulting signal has been processed of multiband compressor dynamics.

For interest we control studio recording the measurement signal chips, which are made in the production logatomic tests in the Czech Radio, they commissioned a dynamic compressor with a compression ratio of 1: 4th The original measured value of the uncoated signal STI = 1.00 (with otherwise the same level corresponding to a level of 70 dB) after dynamic range compression to decrease to STI = 0.91. Speech transmission index fell by almost 10% is, in our opinion, can be considered insignificant, though still moving in the highest quality category A +.

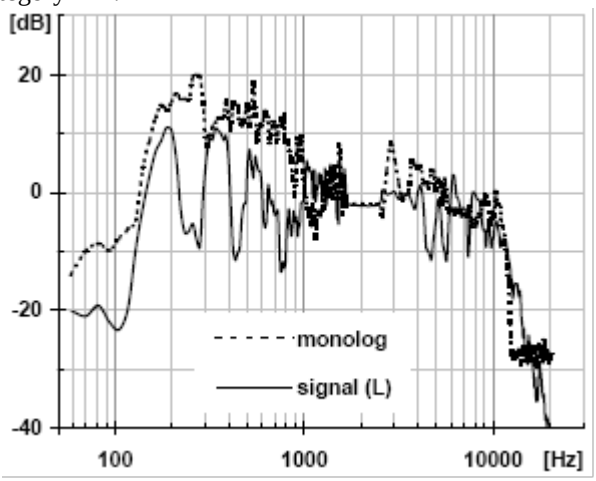

Fig. 8 Comparing the spectra of female voices from logatomic tests and monologues

Figure 8 is a comparison of the spectra of female voices from logatomic tests and corpus monologue. In the subjective assessment of quality of listeners records the recording logatomic tests identified as brighter and more pleasant sound than compressed recording from radio archives.

When comparing the two curves in figure 11 we see a significantly higher level of the compressed signal in the frequency region of 100 Hz to 1 kHz (0 to +10 dB), and much lower levels (0 to -10 dB) at frequencies of 5-10 kHz in sibilants . Voice spectrum logatomic test level-balanced, approximated frequency response is flat and corresponds to adjusting the recording path when shooting tests.

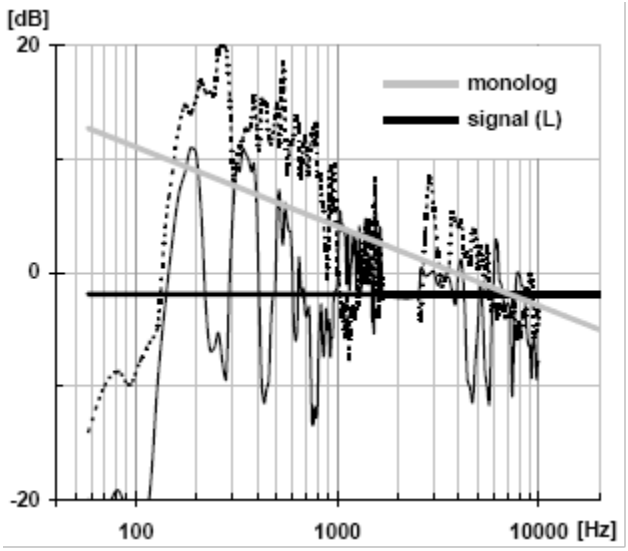

Fig. 9 The approximate frequency characteristics compared records

To approximate the frequency characteristics used frequency range of 150 Hz to 11 kHz, in order to eliminate the steep drop in signal level over a frequency of 12.5 kHz radio recording.

In contrast, the approximated frequency response analyzed the record of monologue corpus shows a -3 dB/oct (-10 dB/dec) and the center of gravity of the signal area compared with the center of gravity has shifted flat frequency response from 4650 Hz to 2929 Hz. The center of gravity of the signal area has only geometric meaning. Its position on the frequency axis can determine the trend of the distribution of signal level of the center frequency band.

X. Conclusion

Based on the measurement results processed so far and part research surveys for the next period we can prioritize each task schedule component activities and verify the selected procedures of possible solutions. It is already clear that the assessment of achievable intelligibility (regardless of the type and destination area) in addition to previously evaluated parameters: reverberation time, frequency characteristics, in connection with a bandwidth of directionality emitters, we will have to consider the needs of educational research and eventual peace dynamic compression audio tracks audiovisual recordings.

References

[1] R. Nemec, S. Hubálovsky, The use systems for measurement using a computer in the school laboratory to explanation for temperatures during the water heating. In: *Education and modern educational technologies (EMET 2014).* Piscataway: IEEE, 2014, p. 103-107.

[2] P. Voborník, Mini-Language for Effective Definition of the Color Gradients. In: *Materials, transportation and environmental engineering II (CMTEE 2014).* Zurich: Trans tech publications, 2014, p. 1882-1885.

[3] P. Voborník, Modification of the perfect cipher for practical use. In: *Manufacturing, engineering, quality and production systems (MEQAPS 2014).* Athens: World scientific and engineering academy and society, 2014, p. 64-68.

[4] S. Hubalovsky, J. Sedivy, Education of student's project team cooperation using virtual communication supported by LMS system. In: *14th International Conference on Interactive Collaborative Learning (ICL2011) - 11th International Conference Virtual University (VU'11)*, Bratislava, Slovenská technická univerzita, 2011, p. 456–459.

[5] S. Hubalovsky Remote desktop access us a method of learning of programming in distance study. In: *14th International Conference on Interactive Collaborative Learning (ICL2011) - 11th International Conference Virtual University (VU'11)*, Bratislava, Slovenská technická univerzita, 2011, p. 450–455.

[6] S. Hubalovsky, Research of Methods of a Multidisciplinary Approach in the Teaching of Algorithm Development and Programming. In: *9th International scientific conference on distance learning in applied informatics (DIVAI 2012)*, Nitra, Univerzita Konštantína Filozofa, 2012. p. 147–156.

[7] R. Nemec, S. Hubalovsky, *Software Design of System SMPSL*. WSEAS transactions on computers. 2014, Vol. 13, p. 329-337. ISSN 2224-2872.

[8] M. Hubalovska, S. Hubalovsky. *Implementation of the Systems Approach in Mathematical Modeling, Dynamic Simulation and Visualization Using MS Excel Spreadsheet*. International Journal of Mathematics and Computers in Simulation. 2013, Vol. 7, No. 2, p. 267-276.

[9] S. Hubalovsky, M. Musilek. *Modeling, Simulation and Visualization of Real Processes in LOGO Programming Language as a Method of Development of Algorithm Thinking and Programming Skills*. International Journal of Mathematics and Computers in Simulation. 2013, Vol. 7, No. 2, p. 144-152.

[10] Hubalovsky, S., Milkova, E.: Modeling of a real situation as a method of the algorithmic thinking development. In: Advanced Educational Technologies, Proceedings of 6th WSEAS/IASME International Conference on Educational Technologies (EDUTE'10), WSEAS Press, Kantoui, Sousse, Tunisia, May 3-6, 2010, pp. 68–72

[11] E. Milková, A. Hůlková, Algorithmic and Logical Thinking Development: base of programming skills. *WSEAS transactions on computers*. 2013, 12(2), s. 41-51.

[12] E. Milková, Development of Algorithmic Thinking and Imagination: base of programming skills. In: *CSCC 2012: proceedings of the 16th WSEAS international conference on communication*. Athens: World scientific and engineering academy and society, 2012, s. 347-352.

Doc. R. Drtina, PhD., was born in 1959 in Czech Republic**.** associate professor on University of Hradec Kralove, Czech Republic. University of Hradec Kralove, Faculty of Education,, Rokitanskeho 62, 500 03 Hradec Kralove, Czech Republic (phone: +420 493331171; e-mail: rene. drtina@uhk.cz). His scientific activities are education and informatics.

Ing. Mgr. J. Sedivy, Ph.D., was born in 1963 in Czech Republic**.** Doctor degree in Theory of technical education in 2006 on University of Hradec Kralove, Faculty of Education, Czech Republic. University of Hradec Kralove, Faculty of Education, Department of Technical Subjects, Rokitanskeho 62, 500 03 Hradec Kralove, Czech Republic (phone: +420 493331171; e-mail: josef. sedivy@uhk.cz). His scientific activities are computer graphics and communications in education and informatics.