jet INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2020, VOL. 66, NO. 1, PP. 139-144
Manuscript received October 25, 2019; revised January, 2020. DOI: 10.24425/ijet.2020.131855

Quality Assessment of Musical and Speech Signals Broadcasted via Single Frequency Network DAB+

Maurycy Kin, Stefan Brachmański

Abstract—**Paper presents the results of quality assessment of speech and music signals transmitted via DAB+ system with the use of Single Frequency Network (SFN). The musical signals were evaluated in overall quality domain. The subjective research was provided with the use of Absolute Category Rating procedure according to the ITU recommendation and the results have been presented as the MOS values for various bit rates. The speech signals were additionally examined with PESQ method. The results have shown that the assumed quality of 4 MOS, for this kind of broadcasting could be achieved at 48 kb/s for speech and 64 kb/s for music. This fact was confirmed by both: subjective and objective research. The comparison between the results obtained for SFN broadcasting with three emitters with single-emitter broadcast was presented.**

Keywords—**quality evaluation, digital broadcasting, single frequency network, speech quality, music quality**

I. INTRODUCTION

IN 2013 European Broadcasting Union recommended an immediate introduction of Digital Audio Broadcasting DAB+ in III VHF band (174 MHz – 230 MHz), with the additional services as programs descriptions, traffic information, the GPS, news, weather reports, city maps and others. It simply causes a growth of signal volume so it is necessary to apply some lossy compression of the audio signals to make all of DAB advantages possible to work [1]. According to the same EBU recommendation from 2013, the DAB system should be applied for large areas (national) as well as for the local communities. For the last purpose, the Single Frequency Network (SFN) seems to be relatively cheaper solution, which enables the local broadcaster to cover small areas, as specific cities and other defined territories. The experimental SFN was built in Wroclaw, Poland, and it is based on the three transmitters making a triangle inner which the receiving of radio programs would be possible with the acceptable quality of experience assumed as good one. The DAB+ standard with HE-AAC v. 2 (*High-Efficiency Advanced Audio Coding*) has been applied for this network what allows to significant increase of a sound quality by the Spectral Band Replication (SBR) as well as Parametric Stereo (PS) processing [2]. As it had been confirmed in previously done experiments [3, 4, 5] the acceptable sound quality, with the SBR and PS being on, can be achieved for 64 kb/s for music, and 48 kb/s for speech.

The synchronization between transmitters is the main problem of this solution because of the disappearing of some sub-carriers causing some lacks of transmitted material as well as the non-periodical decrease of sound quality.

There could be found many aspects of objective and subjective sound quality evaluation of digital sound, especially under lossy compression. Although, the methods and some aspects concentrate usually on the high-order formats of audio [6, 7], they could be applied to sound quality assessment of such media as radio broadcasting, for example. Another interest of quality defining and measurement comes from data reduction of the transmitted signals and it leads to answering the question: how much information could be lost or changed without affecting a subjective quality of the sound? The impairments introduced by a coding system include such deformation of signal as linear distortion, quantization noise, pre-echoes, frequency bandwidth limitations, and changes in timbre and stereo image as well as periodic or non-periodic modulation effects. Another group of original signal distortion results from transmission process (missing sound or extra noise and correlation effects).

The quality of musical as well as speech signal transmission may be assessed with the accordance of 5-degree scale MOS (Mean Opinion Score) corresponding to the range from very annoying to imperceptible distortion, or in a 100-points MUSHRA scale only in musical material case [8]. The test material depends on the method and subject of evaluation: for speech quality assessment the lists of sentences are used [9] while for quality of music the test sequences of various musical material are applied [8, 10].

According to a quality criterion, the assessment is done with the use of Absolute Category Rating (ACR) or Degradation Category Rating (DCR) methods, both recommended by the International Telecommunication Union [8]. In the ACR method listeners are asked to assess the samples using a scale from 1 to 5 in descriptive intervals: from excellent (5), through good (4), fair (3), poor (2) to bad (1), without comparison to the pattern. The DCR method uses also five-point scale but with different meaning: degradation can be inaudible (5), audible but not annoying (4), slightly annoying (3) annoying (2) and very annoying (1). The paradigms of stimuli presentation in both methods are different: in ACR only rated signals are presented to listener while DCR is based on the comparison between the pattern and the sample being evaluated, and the pattern signal is

This work was supported partially by Wroclaw University of Science and Technology, grant no. 049U/0024/19, and National Center of Research and Development, grant no. PBS3/A3/19/2015.

Both Authors are with Wroclaw University of Science and Technology (e-mail: maurycy.kin@pwr.edu.pl, stefan.brachmanski@pwr.edu.pl).

/

always presented as the first in pair. In such application as transmission quality evaluation both methods may be suitable.

For speech signals the objective PESQ (Perceptual Evaluation of Speech Quality) is also recommended for use by ITU [8, 11] and it can be applied as automatic measurement on the base of particular material for transmission monitoring. The idea of PESQ measurement is based on so called internal representation which reflects a theoretical form of speech in human brain. As a reference signal, the previously recorded male and female voices (one sentence by each voice) are used. Such prepared original signal is transmitted via telecommunication channel being under investigation. At the output of this channel the signal is distorted and next, these two signals are compared in a psychoacoustic domain reflecting the human impression of speech. The transformation from the physical signal's form into the psychoacoustic representation is done in three steps: time – frequency reflection, frequency – filtering by the basilar membrane filters (barks or ERB) [7, 12] and scaling of the signal levels. At the end of processing chain, the cognitive model is applied, and the final decision is a result of comparison between two internal (weighted according to the human hearing system) spectra of both: reference and tested signals. The results of PESQ is also expressed in MOS what enables to compare this objective measurement with the subjective one.

In Department of Acoustics and Multimedia at Wroclaw University of Science and Technology, the sound quality assessment of Digital Audio Broadcasting has been investigated since 2005 [3, 4, 5, 13], for one-transmitter mode in that case. Currently made evaluation for SFN is a continuation of this kind of research

II. EXPERIMENT

A. Measurement procedure

Because of the comparison of the results of currently and previously provided research, it was decided to use the ACR method. Moreover, this method reflects a typical, real listening to the radio programs, where the listeners' impressions depend mainly on various aesthetical criteria, without comparison to other sounds. The ACR procedure has been used for both kinds of signals: the musical samples and speech.

B. Test signals

The research was based on the sound material broadcasted by the particular radio stations. Sound samples were recorded at the analog line output of the consumer set DAB Sangean DPR-26 receiver so the signal was passed by the whole transmission channel: multiplex – DAB transmitter – receiver. The ZOOM H4n PRO and TASCAM DR-100 MKIII were used as the recording machines. The samples were transmitted with four different bit-rates (128 kb/s, 96 kb/s, 64 kb/s and 48 kb/s) with sampling frequency of 48 kHz. The possible lowest bit-rate equal to 24 kb/s was not taken into account because of the worse quality offered by this bit-rate [3]. For all the bit-rate values, a Spectral Band Replication (SBR) processor was turned on. The Parametric Stereo processor was applied for the lowest bit-rate (48 kb/s in this experiment) only. As reference signals, CD recordings with the identical two samples to the broadcasted ones, were used.

The sound material was divided into five groups: classical music, rock, jazz, pop and speech. All the samples were taken from various radio programs so it could be said that these signals reflected the typical material broadcasted in the region as well as the whole country. The selected fragments were prepared in accordance with ITU and EBU recommendations [8, 10]. The time of stimuli did not exceed 25 seconds and the musical structures of the samples were taken into account during a sample preparation. The radio programs were recorded in three different places in Wroclaw: at the University of Science and Technology, near Sky Tower (Powstańców Śląskich Str.) and near the city center. The test was constructed in such a way that in the test trial every sample was presented three times in random order and the pattern sample (from CD) was presented one time only due to verification of listeners' notes and the stability of their answers.

For subjective speech assessment, the test lists of Polish sentences were taken also from the real radio programs and prepared in Department of Acoustics and Multimedia, Wroclaw University of Science and Technology [14]. Each list was divided into 10 groups each with 5 tasks. The testing speech material contained the sentences spoken by male and female voices. The listeners assessed coded speech samples and hidden reference signals.

C. Listening condition

The listening tests were performed with the staff of experts containing eleven people (8 male and 3 female), aged from 26 to 33 years old. The listeners had their hearing loss no more than 5dB with the reference to normal hearing, what had been confirmed previously with audiometric tests. The listeners had participated in various listening tests previously and all of them have been working as recording engineers, recording and radio producers. Tests were performed in accordance with EBU and ITU recommendations [8, 10]. The sound material contained various kinds of musical styles as well as speech signals, and all of 13 samples were typical for the profile of National Polish Radio or local radio stations.

The signals were presented with the use of CD-player (Pioneer PD-201) and a pair of active loudspeakers-boxes (TLC Pro-AMS 1). The listening session were conducted in the studio at Department of Acoustics and Multimedia with a noise level less than 35 dBA in four series: each one was 25 minutes long with 20 min. pauses between them. The total time of test run was 2 hours. All the listeners were familiar with the subject of experiment as well as the place of listening. The last fact made the experts the comfortable mood and has eliminated some kind of their deconcentration.

The listeners were asked to give notes in a questionnaire with x-mark at the corresponding note, and after presentation a particular stimuli there was a 6 second break for giving answers.

III. RESULTS AND ANALYSIS

A. Assessment of musical samples

The results of quality assessment of musical signals were treated statistically and the variances of results obtained for the chosen groups of sound material as well as for all the bitrates were found as homogeneous and confirmed by Bartlett test ($\chi^2 = 4{,}17 < \chi_\alpha^2 = 5.98$ at $\alpha = 0.05$). Thus, it allowed to average the results over the eleven listeners in testing group and their three repetitions for four types of the musical program.

The result of statistical treatment by the means of *ANOVA* (ANalysis Of VAriance) test showed that the results could be also averaged over the place of receiving DAB signal, i.e. the results of subjective evaluation did not depend on the investigated regions of the city (F = 1,39 < F_α= 3,68, at α = 0,05). However, the musical style influenced the listeners' opinions (F = 6,53 > F_α = 5,16, at α = 0,05).

The averaged results of the subjective quality of musical signals in dependence of bitrate and for various kinds of music are presented in Fig. 1.

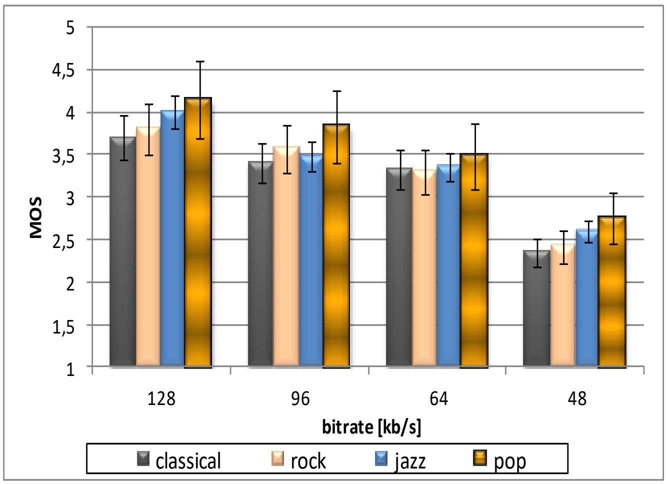

Fig. 1. The averaged MOS results of sound quality for musical signals transmitted via DAB+ system with various bit rates with the range of confidence of 95%.

As it can be seen the MOS values decreases from about 4,0 at 128 kb/s to 2,5 at 48 kb/s bitrate, and this decrease changes monotonically.

When one assumes that the acceptable sound quality can be obtained when the MOS value is 3,0, or higher, the only bitrate of 48 kb/s does not guarantee the satisfactory impression for all musical kinds. It can be also shown that the MOS results for classical music is lower of 0,3 – 0,6 than for other investigated styles. It may be caused by the more critical impressions in aesthetical domain: the sound color as well as the panoramic spread and balance between sound sources should be natural. After the coding process some of the musical instrument may sound heavily synthesized and somewhat distorted.

As it was mentioned above the dependence between sound quality and the bitrate values goes in monotonic manner what enables to estimate the minimal value of bitrate at which the sound quality would be at the assumed level [15, 16]. For example: the lowest bitrate for classical music can be set at 96 kb/s while for pop music the bitrate at which the same sound quality can be achieved is 64 kb/s. It should be noted that in described experiment, the typical sound material for particular radio station profile was used and this material might be very often pre-processed before emission and adapted in dynamic range mainly for the listeners' expectations what it is a typical operation for the commercial radio station or program where the univocal sound is desired. Such pre-processing may introduce some distortion in panoramic spread as well as in the original dynamics and balance [17, 18, 19].

B. *Speech signals assessment*

The quality of speech signals was evaluated, like a musical signals, with ACR method. The statistical treatment with the *t-Student* testing indicated that the listening staff was homogenous (t = 1,67 < t_α = 2,06, at α = 0,05) which enabled averaging the scores over all listeners. Like as for musical signals, with the use of *ANOVA* test the influence of the place of program receiving on the subjective speech quality was not significant (F = 0,87 < F_α = 3,68, at α = 0,05).

In Fig. 2. the results of subjective speech quality examination averaged over all of three places are presented. It can be noticed that there is no great difference between obtained results: MOS values vary from 4,42 to 4,03 at the relatively large range of bitrates from 128 kb/s to 48 kb/s, although these differences are significant (F = 4,52 < F_α = 4,08, at α = 0,05). Also, the MOS values are greater than in a case of musical signals: speech is not as spectrally rich and complicated as a musical performance, in particular the symphonic one. Thus, the decrease of sound quality for music is greater than for speech signals.

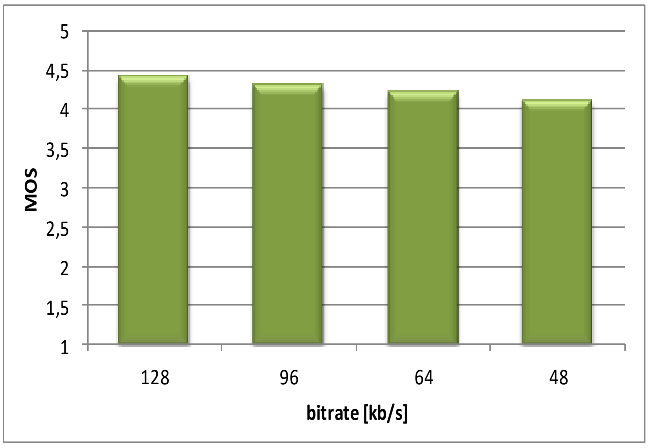

Fig. 2. The averaged results of speech quality assessment in all measurement points in Wrocław.

In order to find the minimal bitrate value which guarantees at least a satisfactory speech quality, and to compare the results of subjective speech evaluation of on-air radio programs the simulation research has been provided. It turned out that the results obtained for HE-AAC format coding with various bitrates gave similar results to these on the base of on-air programs.

The results of that simulation are presented in Fig. 3. as the MOS results obtained for the same listening staff. These results slightly differ in comparison to the on-air material testing (the MOS opinions are greater) and it could be caused by the fact that the sentences' lists had been spoken by the professional speakers (male as well as female) and the pronunciation of speakers' speaking on the radio have not been ideal, so this divergence in results could be explained by this fact. During the testing as well as on-air programs it was authors' duty to prepare the best sounding and pronouncing samples but sometimes, particularly in a case of on-air emission, it was not ideal.

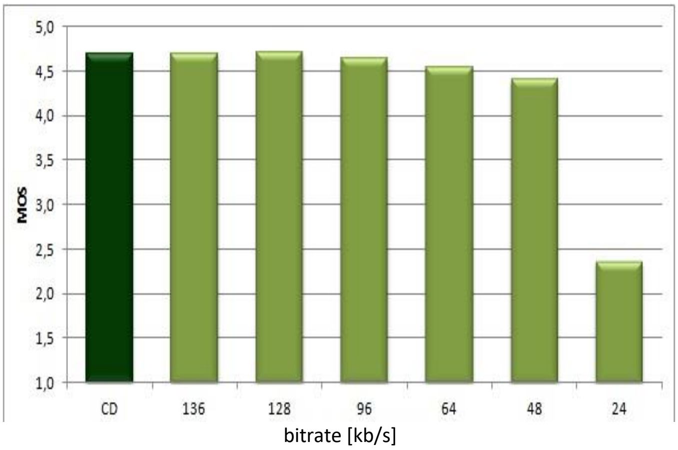

Fig. 3. Averaged subjective evaluation scores for speech signals in a simulation conditions.

On the base of simulation method, the minimal bitrate value which guarantees the acceptable speech quality is 48 kb/s (for the musical transmission, the simulation research gave the results of 64 kb/s, for all kinds of musical styles).

It has been also found that a monotonic relation between the objective measurement MOS-PESQ values and the subjective results obtained by the use of ACR method exists. It should be noted that this fact was confirmed for both: DAB+ signals and for simulation research made for wider range of bitrate values.

Results of objective evaluation obtained with the use of PESQ method and done at the simulation research are presented in the Table I.

TABLE I
THE MOS VALUES OF SUBJECTIVE AND OBJECTIVE PESQ QUALITY ASSESSMENT.

Bitrate [kb/s]	MOS of PESQ	MOS of subjective test
128	4,41	4,42
96	4,39	4,30
64	4,29	4,22
48	4,23	4,03

It can be seen that the MOS values of subjective and objective assessments differ slightly which indicates the good convergence of both methods and these differences may be caused by the cognitive process applied in PESQ while the human senses indicate a different rating when degradation is greater. For example, in PESQ algorithm the loudness loss, delay, some voice effects (resulting of the spectra changes and making the voice sound with special flavors) and other impairments do not affect MOS results significantly, as it has been stated in ITU recommendation [8].

C. Comparison of the sound quality at SFN and the single-transmitter DAB+

As it was mentioned in the introduction, the work on the quality assessment of DAB+ has been continued by the authors for several years, including the experiments with quality evaluation for one-transmitter broadcasting [4, 20]. The results of that research made for musical signal evaluation are presented in Table II.

It can be seen that the values of MOS differ slightly what means that the way of broadcasting (one-emitter mode vs. three emitters in SFN) does not influence the quality assessment of musical signal. Moreover, any artifacts mentioned in the literature [1] for SFN with 3 transmitters, as periodic or non-periodic modulation effects, missing sound or extra noise and signal de-correlation effects have not been observed in any of the two broadcasting modes.

The small differences in MOS values may have their roots in the fact that in previously made evaluation the test signals for broadcasting and evaluation had been prepared as a special trial while in the present research the signals were taken from the actual radio programs, with all the processing effects applied for the achievement of univocal sound for particular radio station, according to its profile and the target of listeners [21]. Such pre-processing of musical material makes the sound usually wider, with boost in a lower – frequency range and sounding as more rich and saturated [4].

In Table III the results of speech quality assessment for single transmitter broadcasting are presented as MOS values for subjective assessment as well as for the PESQ testing.

TABLE III
MOS RESULTS OF SPEECH QUALITY ASSESSMENT FOR ONE-TRANSMITTER BROADCASTING

Bitrate [kb/s]	Subjective test	PESQ
128	4,25	4,82
96	4,19	4,79
64	4,16	4,71
48	4,05	4,65

TABLE II
MOS RESULTS OF SOUND QUALITY ASSESSMENT OBTAINED IN SINGLE FREQUENCY NETWORK AND ONE-TRANSMITTER BROADCASTING

Bitrate [kb/s]	Classical		Rock		Jazz		Pop	
	One-transmitter	SFN	One-transmitter	SFN	One-transmitter	SFN	One-transmitter	SFN
128	3,91	3,78	3,98	3,75	4,01	3,98	4,06	4,11
96	3,64	3,42	3,46	3,52	3,76	3,53	3,95	3,81
64	3,31	3,26	3,40	3,38	3,23	3,31	3,61	3,50
48	2,51	2,35	2,65	2,49	2,81	2,60	2,87	2,75

IV. Discussion

Analyzing the obtained results, it can be stated that at the bitrate of 128 kb/s of the DAB+ system the MOS results are about 4,0 what guarantees the good quality as the good one for both – musical material and speech signals. In the case when the fair quality (with MOS of 3,0 meaning acceptable) is permitted, the bitrate values are different for music and speech. The results have shown that minimal bitrate value is 64 kb/s for music and the 48 kb/s for speech transmission.

It should be noted that the listeners rated the original material of rock music (taken from CD) at the similar level as the DAB+ transmission with 128 kb/s (MOS 4,1) while reference sample with classical music was always assessed with higher MOS value (4,5) in comparison to the broadcasted one with 128 kb/s bitrate. It means that for the assurance of specific transmission with high quality from aesthetical point of view, the higher values of bitrate (128 kb/s, or higher) are recommended. The bitrate of 96 kb/s, or lower may cause the loss of detailed information about a sound atmosphere as well as the spatial sound attributes which are commonly used during the process of classical music recording [4, 24, 25]. The impairments introduced by a coding system may include such deformation of signal as linear distortion, quantization noise, pre-echoes, frequency bandwidth limitations, and changes in timbre and stereo image. However, during the editing process of musical recording, some effects introduced for creating certain aesthetic impressions, for example dynamic compression, are accompanied by additional processing such as equalization, stereo enhancement or reverberations. For this reason, it is still unclear whether average listeners experience the claimed negative effects of signal processing in actual recorded music productions [18]. On the other hand, signal processing with multi-layered sound effects are widely and increasingly applied in popular music mastering in an attempt to boost perceived loudness impression. The higher degrees of compression of dynamics, for example, have been used to make musical pieces to be perceived in many different situations and environments (noisy places, stereo cars, computer loudspeakers or mp3 players) [25] and many daily-musical-consumers do not pay attention for the lower quality of music instead of focusing their attention on emotional aspects of the text, without reference to the pattern.

As it was mentioned above, the changes in sound color are possible to be introduced at the stage of coding process, however the distortion of sound spectra are not so high and the changes in sound color are imperceptible. Moreover, the timbre is the most important factor influencing the subjective evaluation of sound [24] and when the sound spectra changes are hard to perceive and if an audio material does not include a very special stereophonic tricks based on phase-shift processing, the quality of broadcasted sound remains nearly the same. The above described facts mean that when the broadcasted recordings had been prepared with the great amount of spatial attributes (as it is very important feature of classical music recording) the quality of evaluated sounds gets worse in comparison to the original CD quality. These thesis enables to give some directions of sound material preparation by the particular ways of recording, mixing and mastering processes [25].

Taking into account the obtained results of speech quality perception provided with both methods, one can notice that the MOS values of subjective and objective assessment are different but the constant shift between them could be observed. This difference is higher when the signal is more degraded and the MOS of PESQ are higher if the bitrate decreases. However, it has been reported that the intelligibility of words in gramatically-correct, meaningful sentences is higher as twice in comparison to the non-grammatical pseudo-sentences [23, 26] what suggests that some support in listeners' subjective opinion exists increasing the MOS values but the impression of naturalness of speech sounds seems also to be important. On the base of obtained results, it could be said that the degradation of sound color and naturalness at the lower bitrates is not compensated by the meaning-based aspects of speech. Thus, the PESQ MOS notes might be higher in comparison to the subjective opinions because impressions connected to the aesthetical point of view are not taken into a PESQ algorithm [11].

Figure 4 presents a percentage of the quality evaluation at the bit-rate of 64 kb/s with HE-AAC v. 2 coding, found in the ACR procedure and averaged over all kinds of music and over all listeners. It can be seen that this bit-rate guarantees the perception with basic impressions of sound image at the satisfactory level. The other fact which should be taken into account is the fact that listening tests were performed in a recording studio with good acoustical conditions, without any disturbing sounds. These two facts are far away from typical conditions of listening to the radio.

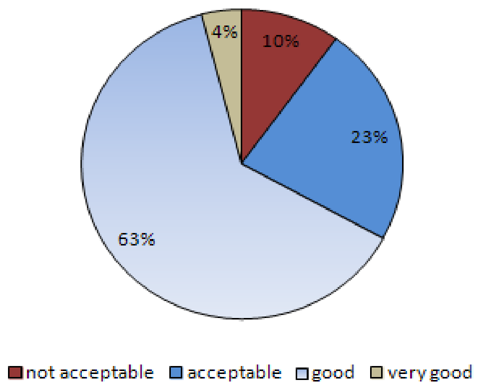

Figure 4. The percentage of ACR assessment at the bitstream of 64 kb/s in HE AAC v.2 coding.

Moreover, the current trend of mixing and mastering, especially of popular music is based on the increase of loudness, in accordance with the rule: better means louder, without taking care of such nuances as timbre, dynamics, space atmosphere etc. [25]. It simply suggests that for some specific groups of listeners and some radio program profiles, DAB+ system should be set with special parameters. For popular and jazz music when the signals are not so complex, the satisfactory sound quality is guaranteed by the bit-rate of 64 kb/s with SBR processing. For more complex sound structures (as a symphonic music, for example), the good quality may be obtained when the bit-stream is of 96 kb/s, or higher.

Analyzing the results of music and speech quality it can be stated that more critical conditions exist for musical samples broadcasting: the evaluation at 64 kb/s resulted in the MOS value of approximately 3,4 for music and 4,2 for speech signals. This difference equal to 0,8 (in an approximation one degree of the quality note) indicates simply that there is more critical conditions which is needed for the music transmission in comparison to the speech broadcasting.

V. Conclusion

On the base of the results of experiments it can be said that the bit-rates higher than 96 kb/s assure the perception with basic impressions of sound image at the satisfactory level for different kinds of music while the speech will be transmitted with the bitrate of 48 kb/s to achieve a good impression and quality.

Comparing the present results with those obtained in the one-transmitter DAB+ it can be stated that the kind of the broadcasting way does not influence the sound quality for both: musical and speech signals. The slight differences may be caused by the fact that in previous experiment the prepared test signals were used while in the present research the sound material contained the pieces or spoken parts taken directly from the actual radio programs.

The obtained results can suggest that the musical recording need to be prepared in a particular way when they have to be reproduced in media introducing the compression of information – this preparation may guarantee the proper transfer of ideas intended by the sound creators. The similar preprocessing was widely applied in classical analogue radio broadcasting (dynamic compression as well as equalization) in order to achieve almost the same sensations when listening to the radio program as to recording. In the present digital days these treatments have not been applied because of the transparency offered of the digital transmission [25]. The need for limitation of information volume affected the return to practicing of sound material initial preparation while it is intended to broadcast via DAB+. The way of such preparation will depend mostly on the target as well as radio station profile, and will take into account the technical parameters of telecommunication channel, the bit-stream value in particular.

It can be also concluded that the objective assessment of sound quality may be applied as automatic monitoring of coding process but the results outcoming from such monitor should be pre-scaled, especially for lower bitrates when speech quality assessment done by objective method results in higher MOS values than in a case of subjective assessment due to aesthetical artifacts introduced mainly by the coding process.

Acknowledgements

This work was inspired by Polish Radio Wroclaw and its Technical Director - Mr. Mirosław Ostrowski who encouraged the authors to the research.

References

[1] W. Hoeg, T. Lauterbach, *Digital Audio Broadcasting, Principles and Applications of Digital Audio*, 2nd ed.; Wiley, England, 2003.

[2] M. Dietz, L. Liljeryd, K. Kjorling, O. Kunz, „Spectral band replication, a novel approach in audio coding". *102nd Convention of Audio Engineering Society, Munich, Germany*. 2002, preprint 5553.

[3] A. Dobrucki, M. Ostrowski, K. Błasiak, M. Kin, S. Maleczek, „Research ofn the sound quality of signals transmitted via DAB+ system". *Przegląd Telekomunikacyjny. Wiadomości Telekomunikacyjne.* R 83 (6): 2010. pp.488 – 491 (in Polish).

[4] M. Kin, „Subjective evaluation of sound quality of musical recordings transmitted via DAB+ system". *134th Convention of Audio Engineering Society, Rome, Italy*, 2013, preprint 8874.

[5] A. Dobrucki, M. Kin, „Subjective and objective evaluation of sound quality of radio programs transmitted via Digital Audio Broadcast (DAB+) system". *Proc. of the ICA, Montreal, Canada*, 2013, vol. 19, doi.org/10.1121/14799170.

[6] P. Kozlowski, A. Dobrucki, „Tuning of the Objective, Perceptual Based Evaluation Methods of Compressed Speech and Audio Signals". *118th Convention of Audio Engineering Society, Barcelona, Spain*, 2005, preprint 6464.

[7] J. G. Beerends, J. A. Stemerdink, „A perceptual speech-quality measure based on a psychoacoustic sound representation". *J. Audio Eng. Soc.*, 42, (3), 1994, pp. 115-123.

[8] ITU-T: Recom, P.800, „Method for subjective determination of transmission quality". Geneva, Switzerland, 1996.

[9] S. Brachmański, „Quality evaluation of speech AAC and HE-AAC coding", *Proc. of Joint Conference Acoustics, Ustka, Poland* [Danvers MA] IEEE, cop. 2018.

[10] EBU Technical Recommendation R22-1999, „Listening Conditions for the Assessment of Sound Programme Material".EBU Geneva, Switzerland.

[11] A. Perry, *Fundamentals of Voice Quality Engineering in Wireless Networks*, Cambridge University Press, 2007, pp. 12-48.

[12] E. Zwicker, H. Fastl, *„Psychoacoustics. Facts and models of hearing"*. Springer, Berlin – Heidelberg, 2007.

[13] S. Brachmański, M. Kin, „Quality evaluation of sound broadcasted via DAB+ system based on a single frequency network". *144th Convention of Audio Engineering Society, Milan, Italy.* preprint 10004, 2018.

[14] S. Brachmański, *Selected topics of the methods of quality assessment of speech transmission,* Oficyna Wydawnicza Politechniki Wrocławskiej we Wrocławiu, 2015 (in Polish).

[15] P. Gilski, Stefański J., „Subjective and objective comparative study of DAB+ broadcast system". *Archives of Acoustics,* 42 (1), 2017, pp. 3 - 11.

[16] P. Gilski, „DAB vs DAB+ Radio Broadcasting: a Subjective Comparative Study". *Archives of Acoustics*, 42, (4), 2017, pp. 715 – 723.

[17] B. De Man, K. McNally, J. D. Reiss, „Perceptual evaluation and analysis of reverberation in multitrack music production". *J. Audio Eng. Soc.*, 65, (1/2), 2017, pp.108-116.

[18] J. Hjortkjaer; M. Walter-Hansen, „Perceptual effects of dynamic range compression in popular music recordings.", *J. Audio Eng. Soc.*, 62, (1/2), 2014, pp. 37-41.

[19] R. Taylor, W. Martens, „Hyper-compression in music production: listener preferences on dynamic range reduction." *136th Convention of Audio Engineering Society, Berlin, Germany*, 2014, preprint 9022.

[20] S. Brachmański, M. Kin, „Assessment of speech quality in Digital Audio Broadcasting (DAB+) system". *134th Convention of Audio Engineering Society, Rome, Italy*, 2013, preprint 8829.

[21] I. McGregor, S. Cunningham, „Comparative Evaluation Of Radio And Audio Logo Sound Design". *J. Audio Eng. Soc.*, 63, (11), 2015, pp. 876-887.

[22] J. L. Flanagan, *Speech analysis, synthesis and perception*. Springer Verlag, Berlin, 1965.

[23] G. A. Miller, „Decision units in the perception of speech". *I. R. E. Trans. Of Inform. Theory*, IT-8, 1962, pp. 81-83.

[24] D. Ko, W. Woszczyk, „Virtual acoustics for musicians: subjective evaluation of a virtual acoustic system in performance of string quartets". *J. Audio Eng. Soc.*, 66, (9), 2018, pp.712-723.

[25] F. Rumsey, „Recording In The Light Of New Technology". *J. Audio Eng. Soc.*, 63, (12), 2015, pp. 1053-1057.

[26] J. G. A. Barbedo, A. Lopez, „A new cognitive model for objective assessment of audio quality". *J. Audio Eng. Soc.*, 53, (1-2), 2005, pp. 22-31.