

Subjective Assessment of the Speech Signal Quality Broadcasted by Local Digital Radio in Selected Locations in Wrocław under Studio and Home Conditions

Stefan Brachmański, Maurycy Kin, and Patrycja Zemankiewicz

Abstract—In October 2018, local digital radio was launched to cover the agglomeration of Wrocław. The implementation of this undertaking required many tests, including qualitative ones, that refer to both music and speech. This paper presents the results of subjective tests based on the evaluation of speech quality of signals recorded at various points in Wrocław. Measurements were carried out in accordance with the recommendations of the International Telecommunication Union as well as in ordinary acoustic conditions in listeners' flats. The rating was made for male and female voices. The most important conclusion is that for speech signal assessment in meaning of the quality the test conditions do not influence the obtained results. The other fact confirmed in the experiment was that the receiving place of DAB+ signal in the Single-Frequency Network also does not affect the perceived voice quality

Keywords— Digital Audio Broadcasting; speech quality; quality assessment

I. INTRODUCTION

THE regular broadcasts of Digital Audio Broadcasting (DAB+) in Poland dates on October 1st, 2013. According to the recommendations of the European Broadcasting Union (EBU), digital radio coverage should cover both large areas (the entire region/country) and smaller, for example urban agglomerations [1,2].

On January 19, 2018, a local DAB+ multiplex was launched in Wrocław, developed as part of a project carried out by the Wrocław University of Science and Technology, the Communications Institute in Wrocław and Polish Radio Wrocław. The solution was implemented as a Single Frequency Network (SFN) in the DAB+ digital radio system. The main advantage of this type of network is the much more efficient use of spectral space. A network of multiple transmitters operating at the same frequency, transmitting the same signals and meeting the relevant synchronization requirements can cover a much larger area and fill the gaps in the range (gap filling technique), while minimizing the total radiated power [3].

The implemented network is based on three transmitters located on three buildings of: Polish Radio Wrocław S.A., the Institute of Telecommunications in Wrocław and the premises of Municipal Water and Sewage Company (MPWiK). The last place is situated near the campus of the Wrocław

University of Science and Technology (Fig. 1). Such localizations of the emitters guarantee the good coverage of the almost entire Wrocław agglomeration [3].



Fig. 1. Localization of DAB+ transmitters in Wrocław

Standard HE-AAC (High-Efficiency Advanced Audio Coding) v.2 [4,5] was used to encode the signals, thanks to which it was possible to improve the quality at lower bit rates using the following processing techniques: Spectral Band Replication (SBR) and Parametric Stereo (PS) [6], allowing a significant improvement in sound quality at speeds of 64 kbps and 48 kbps [2, 7-10].

Measurement as well as an evaluation of the quality of speech and music signals transmitted on DAB+ digital radio has been the subject of many publications [11-14]. These works presented the results of tests conducted under laboratory conditions, in accordance with ITU recommendations, or under standardized conditions used in telephometric measurements [8,9]. The above conditions, however necessary for testing individual attributes or parameters of audio signals, do not reflect the actual conditions of perception of radio broadcasts. Therefore, the authors decided to conduct speech signal quality evaluation studies under laboratory conditions as well as under actual listening conditions of the test material at the listeners' residence, using home listening equipments.

The main aim of the study was to perform speech signal quality tests under laboratory conditions, as well as under home listening conditions. In addition, it was checked whether

Authors are with Wrocław University of Science and Technology (e-mail: stefan.brachmanski@pwr.edu.pl, maurycy.kin@pwr.edu.pl, zemankiewicz.patrycja@gmail.com).



the evaluations obtained under the two conditions differ significantly, and whether the conditions for receiving radio signals in the DAB+ single-frequency network system previously simulated and then measured [3] affect the results of evaluating the quality of transmitted signals. An additional result of the study was to determine the minimum bit rate value that guarantees good quality of speech signals from the listeners' point of view.

II. RESEARCH METHOD

A. Choice of procedure

Speech quality was measured using the qualitative criterion of the Absolute Category Rating (ACR) method recommended by the International Telecommunication Union [15]. The quality of speech is assessed by the listener using a five-point rating scale without comparison with the pattern signals what means that the individual patterns for listeners had been created in their brains. The rating scale reflects the audience's impressions from the best quality, with a value of 5, to the worst, expressed as – 1 [16].

The MUSHRA procedure with continuous scale from 0 to 100 has been also considered for the use but this methodology has been used mainly for the assessment of low bitrate codecs [17-19].

Listening tests were carried out for two conditions:

1. in the listening studio of the Department of Acoustics, Multimedia and Signal Processing, Wrocław University of Science and Technology,
2. at the home of the listeners.

The measurement for the 1st condition, i.e., in the listening studio, was performed in a room that fulfills the ITU-T P.800 recommendations [15] for the test environment. Its additional advantage was that all participants of the research knew about it, so it did not affect the distraction of the audience related to adaptation to the research site. Each listener, after listening to the teaching sequence (training), and then after listening to the assessed list of sentences, had to give a grade on a scale of 1 to 5 that corresponded to their feelings.

The measurement in 2nd condition, i.e. in the common living spaces, was also performed using the technique described in the ITU-T P.800 document [15], but the electro-acoustic parameters of the room in which the listening session was performed did not necessarily meet the requirements of this recommendation. As in the first case, the room did not distract the listener because it was familiar to him. In addition, the listening session was carried out with the use of electroacoustic equipment daily used by the listeners.

B. Preparation of test material

The test material consisted of 10 sentence lists recorded by a woman and a man with a sampling frequency of 44,1 kHz and a resolution of 16 bits [1,3]. The selection of the sampling rate was determined by the standards adopted in the LocalDAB radio system in the preparation of the broadcast material.

The test lists were broadcast at five-bit rates: 32 kbps, 48 kbps, 64 kbps, 96 kbps and 112 kbps. Each radio station was assigned a one-bit rate, which means that five radio channels were broadcasting test sets instead of normal programs. The test signals received by the digital radio receiver Sangean DPR-26

DAB were recorded on the digital recorder ZOOM H4n PRO [20].

The measuring points (recordings) were established as follows. The focal point was the C-5 building of the Wrocław University of Science and Technology. Two circles were drawn around the central point, the first with a radius of 3.5 km and the second with a radius of 7 km. In these circles, measuring points spaced every 60° were marked. The arrangement of the measurement points is shown in Fig. 2. Measurement points for which a subjective quality assessment was performed are marked with red circles, while blue stars on the outer circle - points for which no measurements were taken. In total, measurements were made for 9 points, including the central point, 6 on the inner circle, one on the outer circle and one close to the outer circle (point no. 6). These measurement points were selected in order to verify that the sites modeled and checked for propagation [3] (different distances from the transmitters, natural obstacles) were not burdened with any disturbances, sometimes occurring during receiving digital signals (interferences, unwanted echoes, etc.), degrading the quality of perceived speech signals. An additional reason for choosing these points was that they were located in the area of the largest human concentrations in the city.

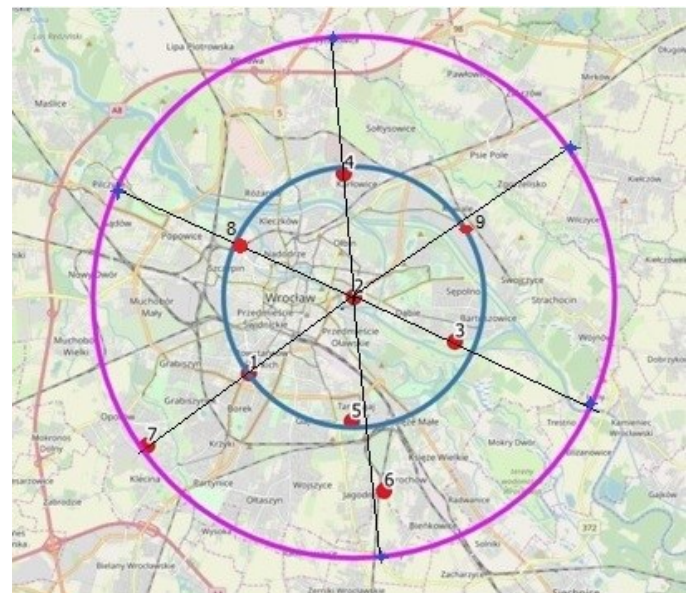


Fig. 2. Distribution of measuring points in Wrocław

C. Implementation of measurements

The listening team was selected from among ordinary listeners, according to the ITU-T P.800 recommendation [15]. In both types of measurements, i.e. laboratory and home listening, the listening team consisted of 30 listeners ranging in age from 18 to 30 years. The test material was previously made available to participants for measurements at home in the form of wave files. The files are numbered in the order in which they are to be evaluated. To control the time, the files have been grouped into folders so that the listening time of all files in the folder being listened to does not exceed 20 minutes. After listening to the track, a break should be taken. In addition to the audio files, an MS Excel sheet with the evaluation form and detailed instructions on how to perform the measurement has been sent to the listening staff.

Before starting the measurements, the audience has been informed about the measurement procedure, including the method of assessing the sentences. The audience gave their grade after listening to the group of sentences. Each group consisted of 5 sentences. The duration of one group of sentences was about 10 seconds. After each group of sentences, there was an 8-second silence during which the listeners assessed the quality of the group they heard on a five-point scale. After averaging over the listeners, for each measurement point and each transmission condition, the score was reported as the MOS (Mean Opinion Score).

The test sentence lists were played back under laboratory conditions with the use of a digital recorder, on which lists of the examined points were recorded. The test signals were presented to the audience through a TLC AMS 1 loudspeaker set with a flat frequency response in the range of 50 Hz - 20 kHz ± 1 dB. The test signals were presented at the level of about 75 dBA. The acoustic conditions of this room met the standards included in the ITU Recommendation P.800 [15].

In the case of home listening, the signals were reproduced using a computer and the listening session was performed with headphones or with the use of a loudspeaker set. The choice of the listening method was done by the listener based on his preferences. The purpose of this procedure was to create listening conditions for the listener when listening to spoken and music programs every day.

The duration of one session did not exceed 2 hours, with a 10 - minutes break after every 20 minutes of listening. The duration of one sentence list was approx. 2 minutes.

Before starting the measurement session at home, participants received instructions from the person conducting the measurements via videoconference. After the training,

the participants performed the test independently, in their own homes. In case of any problems, they had the opportunity to contact the researchers. Completed MS Excel spreadsheets were sent by email.

III. RESULTS AND DISCUSSION

A. Listening in the studio

The average assessment results obtained for nine points of receiving, separately for the male and female voices, are presented in Fig. 3 and Fig. 4, respectively.

When analyzing the results obtained at various points in Wroclaw, slight differences in the ratings for a given bit rate can be noticed. The biggest differences are for the bit rate of 32 kbps, especially for male speaker.

Using the ANOVA test at the significance level of $\alpha = 0.05$, it was shown that the location where the radio signal did not affect the evaluation of the quality of the speech signal for both: the male and female voices ($F = 0.015$ and $F = 0.097$, respectively, with the $F_{\alpha} = 2.208$). These results allowed to average the scores in different parts of Wroclaw for each speed. The subjective evaluation results averaged based on measurements made at nine points, for all bit rates are presented in Fig. 5.

When assessing the results obtained for speech signals, it can be concluded that for the bit rate of 32 kbps the quality was not satisfactory to the listeners. For the bit rate currently used in LocalDAB Wroclaw, which is 112 kbps, the MOS index is at the level of 4.45, which corresponds to the quality of the original signal.

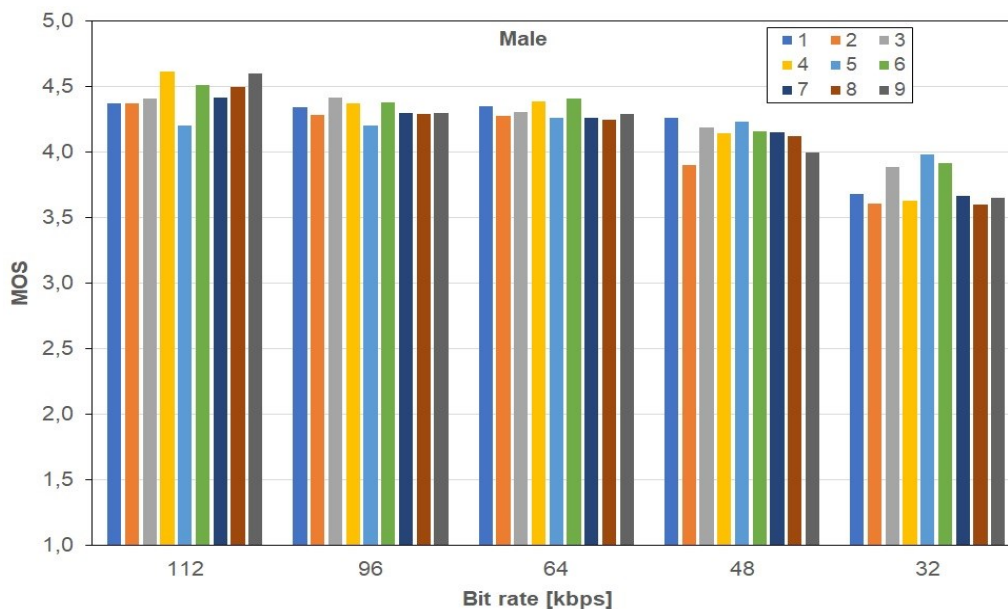


Fig. 3. Assessment of the MOS quality for the male voice at 9 points in Wroclaw for 5 speeds - measurement in the studio

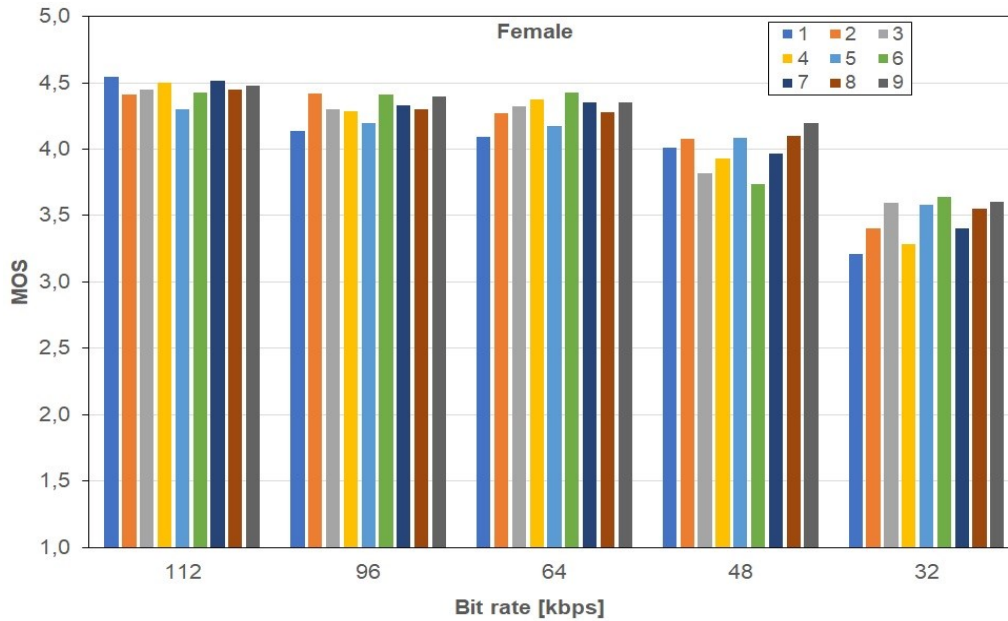


Fig. 4. Assessment of the MOS quality for the female voice at 9 points in Wrocław for 5 speeds - measurement in the studio

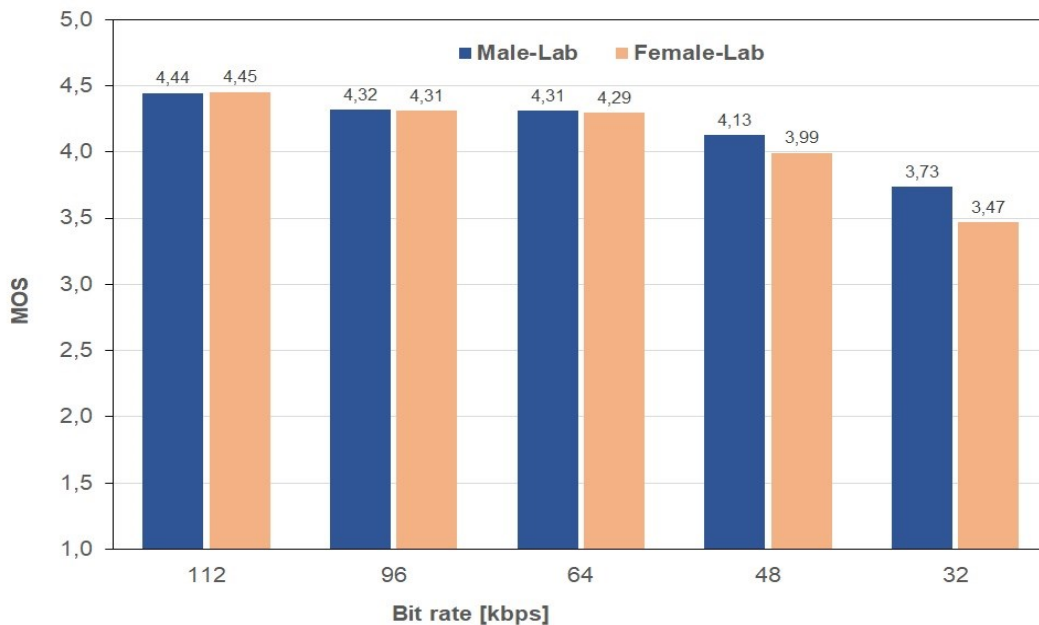


Fig. 5. Speech quality assessment results averaged over all examined points in Wrocław - measurement in the studio

According to ITU-T Recommendation P.800 good quality speech signal transmission features the MOS = 4.0. On the base of the obtained results, it can be concluded that for bit rates of 48 kbps, or higher, a good quality of speech signal transmission was achieved for all the receiving places. For the bit rate of 32 kbps the good speech quality had not been achieved in any receiving points. It means that for this bit rate value the speech transmission cannot be used for commercial broadcasting, but only for special applications

Using the Student's t -test, it was checked whether the differences between the averaged assessment for male and female voice quality were statistically significant. It has been found that there is no reason to reject the hypothesis of insignificance of differences between the quality assessments for male and female voices made at the studio conditions

($t = 0.09 < t_{\alpha} = 2.13$, at $\alpha = 0.05$). The averaged results will be shown in Fig. 7, together with the results of home measurements.

B. Home listening

The same test material was explored for home listening. The files were sent to listeners via email as wave files and the procedure was in accordance with the technique described in the ITU-T P.800 recommendation, but the electro-acoustic parameters of the room in which the listening session was performed did not necessarily meet the requirements of this recommendation. This part of experiment was made for the check if the normal, home listening of the radio programs influences the quality of speech signals.

The significance of differences in the scores obtained was checked by the ANOVA test. It turned out that at the significance level of $\alpha = 0.05$, there was no influence of the place of receiving the radio signal on the evaluation of the quality of the speech signal for the male and female voices ($F = 0.3239 < F_{\alpha} = 3.885$, and $F = 0.056 < F_{\alpha} = 3.885$, respectively). These results allowed to average the ratings of the listeners for each speed, separately for the male and female voices. The result of this averaging is shown in Fig. 6.

Using the Student's t -test, it was checked whether the differences between the averaged assessment for male and female voice quality were statistically significant. It has been found that the differences between the quality assessments for male and female voices made at home conditions are insignificant ($t = 0.01 < t_{\alpha} = 2.13$, at $\alpha = 0.05$). This result allows for averaging the quality of male and female voices, as it has been done for studio listening.

C. Comparison of the results of studio and home measurements

The result of averaging the evaluation of the speech quality obtained under listening conditions in the studio and at home, as a function of the tested bit rates is presented in Fig. 7.

When comparing the averaged results of the speech quality assessments obtained in the listening studio and at home, one can notice slightly higher values of the ratings obtained at home. Higher ratings given by listeners in-home measurements are probably caused by the everyday equipment they used and habits to its sound. Also noteworthy are the high-quality ratings obtained at relatively low bit rates compared to the quality assessment of music signals, where not only the overall sound quality is assessed, but also its individual attributes [13]. The high MOS values may also be due to the more tolerant criterion for assessing speech quality in comparison to the music signals in addition to the lack of interference in speech intelligibility, which undoubtedly affects the evaluation.

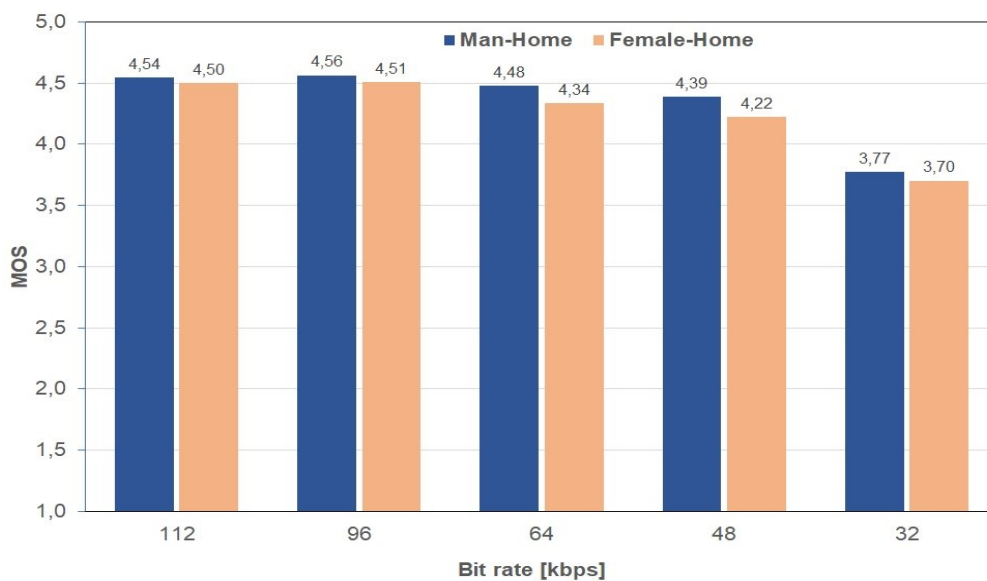


Fig. 6. Results of speech quality assessment averaged over the examined points in Wroclaw - home measurement

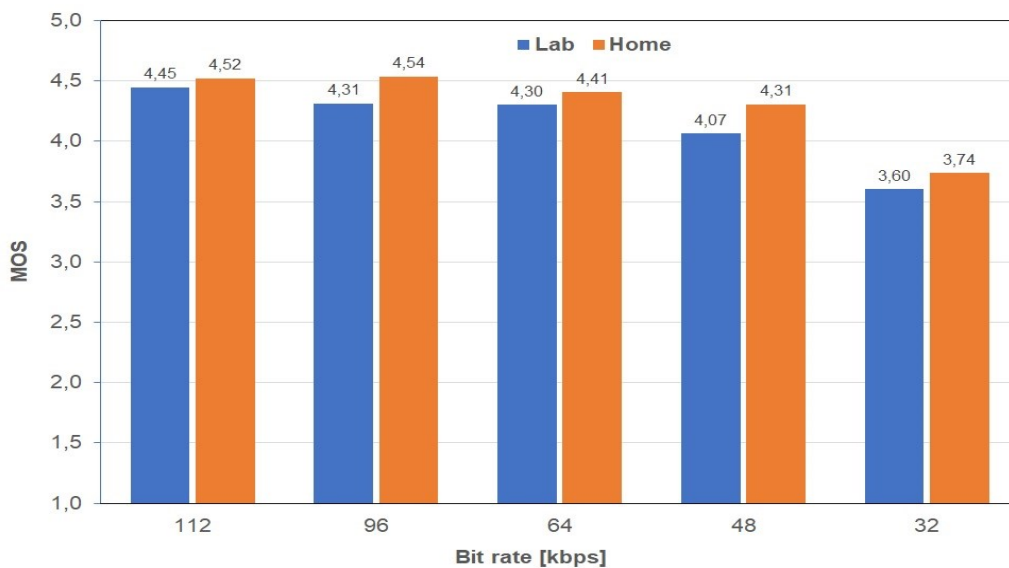


Fig. 7. Averaged results of speech quality assessment obtained in the laboratory and home conditions

Due to the small differences in quality assessment in studio and home listening conditions, it was decided to check the statistical reliability of the obtained assessments. The significance of the differences was checked with the t -test. The result of the t -test showed that at the significance level of $\alpha = 0.05$ that there is no reason to reject the hypothesis of the insignificance of differences between the MOS results of speech quality made in the listening studio and at home ($t = 0.004 < t_{\alpha} = 2, 13$). This can be a major simplification in the organization of research on the evaluation of the quality of speech signals intended for commercial broadcasting; then the study of overall quality is not associated with a specific, well-defined place with very strongly defined acoustic conditions. An important role is also played by the lowering of quality requirements, especially for radio voice, and so-called radio sound in general. This lowering, resulting from the use of an increasing variety of media, is caused by the need to understand the content rather than aesthetic expectations [21].

On the basis of the fact that the results obtained for HE-AAC format coding with various bitrates gave similar results of speech quality to those on the base of on-air programs both in the case of single transmitter emissions and in the case of SFN [13]. Thus, it can be concluded that the preparation of test material as well as the performance of speech signal quality assessment tests can be carried out not necessarily on-air. Also based on previous work [7, 10, 12], it can be predicted that while meeting good quality for speech transmission for a bit rate of 48 kbps, or higher, the good quality for musical signals can be achieved at a bitrate of 64 kbps and HE-AAC coding, without the need for time-consuming and expensive listening experiments.

CONCLUSION

The analysis of the results obtained showed that, although there are slight differences in the assessment of the quality of the speech signal obtained as a result of laboratory measurements (listening studio) and in-home conditions, these differences are insignificant. Higher ratings were obtained during measurements at home, which seems to be caused by the listener's habituation to the sound of his electroacoustic equipment. The smallest difference in the assessment of the quality of the speech signal (Δ MOS) was obtained for the speed of 112 kbps (Δ MOS = 0.07), and the largest for the bit rates of 32 kbps (Δ MOS = 0.25) and 96 kbps (Δ MOS = 0.23). Statistical analysis based on the t -test showed that the obtained differences are statistically insignificant.

The research also showed that in selected parts of Wrocław, the differences in laboratory assessment of the quality of the speech signal of male and female voices are statistically insignificant. Home research has provided similar conclusions. Therefore, it can be assumed that the quality of the speech signal does not depend on the selected sites of signal receiving: in Wrocław located within the districts with a radius of 3.5 km and 7 km (2 points) set in relation to the central point. This means that there is no degradation of emission parameters at the designated points, according to the simulations and measurements which had been made earlier for radio DAB+ signals [22].

According to the ITU-T P.800 recommendation, the speech signal transmission quality can be considered as good at

MOS - 4.0. The obtained results of laboratory and home measurements indicated that at minimum bit rate value of 48 kbps the good quality of speech can be achieved. It means that in order to ensure good quality of speech signal transmission (MOS \geq 4.0) the bit rate should be at least 48 kbps. When analyzing the obtained results, it can also be noticed that starting from 64 kbps, no differences in the evaluation of the quality of the speech signal are observed. It allows the use of variable bit rates of transmitted programs depending on the content contained [12, 13].

The analysis of the results obtained confirms the conclusions obtained in simulation research based on AAC coding [23], as well as on the evaluation of live broadcasts [9]. The previous research showed that the MOS value above 4 is already obtained at a bit rate of 48 kbps. It should also be mentioned that the results obtained for SFN emission differ slightly from the DAB+ signal using one transmitter [7, 10]. Therefore, it can be concluded that the method of emission (single-frequency network vs. one transmitter) does not significantly affect the assessment of the sound quality transmitted in the DAB+ system.

REFERENCES

- [1] D. Niewiadomski, R. Michniewicz, J. Sobolewski, D. Więcek, „Planning of the local Single-Frequency-Network of DAB+” („Planowanie lokalnej sieci jednoczesnościowej radiofonii cyfrowej DAB+”), *Przegląd Telekomunikacyjny - Wiadomości Telekomunikacyjne*, 6, 513 – 516, 2017 (in Polish). <https://doi.org/10.15199/59.2017.6.83>
- [2] W. Hoegh, T. Lauterbach, „Digital Audio Broadcasting: Principles and Applications of DAB, DAB+ and DMB”, 3rd ed., Wiley, England, 2009. <https://doi.org/10.1002/9780470746202>
- [3] I. Michalski, R. J. Zieliński, „Verification Methodology of Fade Characteristics in a DAB+ SFN in Wrocław”, *Int. Journal of Electronics and Telecommunications*, vol. 67, no 3, 537-542, 2021. <https://doi.org/10.24425/ijet.2021.137844>
- [4] K. Brandenburg, “MP3 and AAC explained”, in: *Proc. of 17th International Conference on High Quality Audio Coding*, Audio Eng. Society, Florence, Italy, 1999.
- [5] K. Sayood, “Introduction to Data Compression”, 4th Edition, Elsevier, Waltham, MA, 2012.
- [6] J. Breebaart, S. Van de Par, A. Kohlrausch, E. Schuijers, “Parametric Coding of Stereo Audio”, *EURASIP Journal on Applied Signal Processing*, 9, 1305-1322, 2005.
- [7] S. Brachmański, M. Kin, “Quality evaluation of sound broadcasted via DAB+ system based on a single frequency network”, in: *Proc. of 144th Convention of Audio Eng. Society*, Milan, Italy, preprint 10004, 2018.
- [8] S. Brachmański, “Test material used to assess speech quality in Poland”, in: *Acoustics, acoustoelectronics and electrical engineering*, F. Witos (ed.), Gliwice, 65-79, 2021.
- [9] P. Gilski, “DAB vs DAB+ radio broadcasting: a subjective comparative study”, *Archives of Acoustics*, 42 (4), 715–723, 2017. <https://doi.org/10.1515/aoa-2017-0074>
- [10] M. Kin, “Subjective evaluation of sound quality of musical recording transmitted via DAB+ system”, in: *Proc. of 134th Convention of Audio Eng. Society*, Rome, Italy, preprint 8874, 2013.
- [11] A. Dobrucki, M. Kin, “Subjective and Objective evaluation of sound quality of radio programs transmitted via Digital Audio Broadcast System[DAB+]”, in: *Proc. of Meetings on Acoustics*, 19, 1, 1-7. 2013.
- [12] P. Gilski, J. Stefański, “Subjective and Objective Comparative Study of DAB+ Broadcast System”, *Archives of Acoustics*, 42 (1), 3-11, 2017. <https://doi.org/10.1515/aoa-2017-0001>
- [13] M. Kin, S. Brachmański, “Quality assessment of musical and speech signals broadcasted via single frequency network DAB”, *Int. Journal of Electronics and Telecommunications*, vol. 66, no 1, 139-144, 2020. <https://doi.org/10.24425/ijet.2020.131855>
- [14] P. Počta, J. G. Beerends, “Subjective and Objective Assessment of Perceived Audio Quality of Current Digital Audio Broadcasting Systems and Webcasting Applications”, *IEEE Transaction and Broadcasting*, 61, 3, 407-415, 2015. <https://doi.org/10.1109/TBC.2015.2424373>

- [15] ITU-T Recommendation P.800, Methods for subjective determination of transmission quality, Geneva, Switzerland, 1996.
- [16] S. Brachmański, "Selected problems of speech transmission quality assessment" ("Wybrane zagadnienia oceny jakości transmisji sygnału mowy"), Wrocław University of Science and Technology Edition, 2015, (in Polish).
- [17] ITU Recommendation BS.1534-1, Method for the subjective assessment of intermediate quality level of coding systems, Geneva, Switzerland, 2003.
- [18] A. Perry, "Fundamentals of Voice-Quality Engineering in Wireless Networks, Cambridge University Press, Cambridge, 2007.
- [19] S. Zieliński, "Is a Multi-Slider Interface Layout Responsible for a Stimulus Spacing Bias in the MUSHRA Test?", *Archives of Acoustics*, 40, 4, 585-594, 2015. <https://doi.org/10.1109/TBC.2015.2424373>
- [20] N. Rurzyńska, "Realization of acoustical signals database to assess of speech quality in digital radio LocalDAB+ in Wrocław", ("Realizacja akustycznej bazy sygnałów testowych do oceny jakości mowy w radiofonii cyfrowej LokalDAB na terenie Wrocławia"), Wrocław University of Science and Technology, Engineering Thesis, 2019 (in Polish).
- [21] F. Rumsey, "Telling the difference – preference and prediction", *J. Audio Eng. Soc.*, 68, (10): 774 -778, 2020.
- [22] R. J. Zieliński, "Conditions for obtaining correct DAB+ signal reception on a single frequency network", („Warunki uzyskania poprawnego odbioru sygnału DAB+ w sieci jedno-częstotliwościowej”), *Przegląd Telekomunikacyjny - Wiadomości Telekomunikacyjne*, 6, 509 – 512 (in Polish). <https://doi.org/10.15199/59.2017.6.82>
- [23] S. Brachmański, "Quality evaluation of speech AAC and HE-AAC coding", in: *2018 Joint Conference – Acoustics*, Ustka, Poland; 2018. <https://doi.org/10.1109/ACOUSTICS.2018.8502335>