Why Do Older Adults Prefer Some Radio Stations? Helping to Increase Speech Understanding

Artur Janicki and Krzysztof Szczypiorski
Institute of Telecommunications, Warsaw University of Technology, 15/19 Nowowiejska Str., 00-665 Warsaw, Poland
Email: {a.janicki, ksz}@tele.pw.edu.pl

Abstract —This paper addresses the problem of older adults' understanding of speech. By analyzing speech rates used by various radio stations and by confronting them with audience ratings we found that there is some correlation between speech rate used by radio stations and audience preferences in different age groups. We present various studies analyzing the problem of speech perception among older adults; we also show speech processing algorithms, which are able to modify speech rates. In the experimental section of the study, we run preliminary intelligibility tests using original, fast speech, and slowed-down speech using a time-domain prosody modification algorithm. The results suggest that older adults can benefit from a slowed-down rate of speech.

Index Terms—Elderly, speech perception, radio broadcasting, prosody, TD-PSOLA, assistive technology

I. INTRODUCTION

Our societies are growing older due to lower mortality and lower fertility; population aging is taking place in nearly all countries of the world. The global share of older people (understood as aged 60 years or more) increased from 9.2% in 1990 to 11.7% in 2013; this number is expected to reach 21.1% by 2050 [1]. What is more, the number of people reaching 80 years or more is increasing. It is estimated that if current tendencies are preserved, by 2050 there will be 392 million people aged 80 years or older, i.e. three times more than now [1].

While lower mortality and a longer life are undoubtedly positive things, a growing number of older adults are a challenge for society. One of the aspects that needs to be addressed is the problem of assuring a decent life of comfort for older adults, despite various deficiencies typical of older people.

One of the problems inherent to older people is that of decreased understanding of speech. Age typically affects speech perception – it is estimated that 30% of people over 65 years should use a hearing aid [2], as their hearing impairment justifies the use of such a device. In our article, we address the problem of speech perception among older adults in a bit of a peculiar way – by analyzing, which radio stations older adults listen to. We will also try to suggest a way to increase speech understanding among older adults using speech processing methods.

Manuscript received June 14, 2015; revised November 18, 2015.
Corresponding author email: a.janicki@tele.pw.edu.pl.
doi:10.12720/jcm.10.11.926-931

TABLE I: AUDIENCE RATINGS (IN PERCENTAGES) FOR WARSAW, POLAND, FOR VARIOUS AGE GROUPS FOR SIX RADIO STATIONS MOST POPULAR AMONG OLDER ADULT LISTENERS (BASED ON [3])

Radio station / Age [yrs]	15-24	25-44	45-64	65-75	All
# Participants	706	2162	1635	563	5066
PR1	2.3	3.6	12.0	35.4	10.3
PR3	6.1	10.5	14.4	5.6	11.0
RM	0.0	0.7	3.3	10.6	2.7
RMFFM	11.0	10.3	9.5	5.2	9.5
TOKFM	1.4	4.5	8.1	17.2	7.0
ZET	11.0	13.6	17.1	6.6	13.8
Total	31.8	43.2	64.4	80.6	54.3

A. Motivation

It is commonly known that older radio listeners prefer different radio stations than the young and middle-age audience. To learn more details of this difference, the authors analyzed radio audience ratings, realized by the Polish National Broadcasting Council (Krajowa Rada Radiofonii i Telewizji, KKRiT) [3], which queried 5,066 (2,654 female and 2,412 male) radio listeners in Warsaw, Poland in 2012. Table I shows the results of audience rating for various age groups for the six radio stations most popular among older listeners in Warsaw at that time. Audience ratings were calculated according to the following formula:

$$AR_s[\%] = \frac{\sum_{i=1}^{N} t_s(i)}{\sum_{i=1}^{N} t(i)} * 100\% \qquad (1)$$

where AR_s denotes audience rating for a radio station s, $t_s(i)$ denotes the amount of time listener i listens to the radio stations, $t(i)$ denotes the amount of time listener i listens to the radio in general, N denotes size of the age group.

As one can observe, there are significant differences among the age groups. While the presented six radio stations (denoted with acronyms: PR1, PR3, RM, RMFFM, TOKFM, ZET) constitute more than 80% of the time that older adults spend listening to the radio, they cover less than 32% of such a time for the youngest listeners (aged 15-24 years). Older adults prefer mostly PR1, TOKFM and RM (in order of preference); in contrast, listeners younger than 45 years of age prefer ZET, RMFFM and PR3.

There might be multiple reasons for the popularity (or unpopularity) of certain radio stations among listeners from a given age group. Let us just mention some of them:

- Character of a station (universal, news, musical, religious, business, etc.);
- Music played (“adult contemporary”, rock, jazz, oldies, chill-out, classical, etc. [3]);
- Ratio between oral and musical content;
- Amount of commercials played;
- The way how people speak on air.

In this work, we will concentrate on the last aspect. In particular, we will try to address the issue of the rate of speech at which people speak when talking to the audience on the radio; we suspect that it can be related to the preferences of elderly radio listeners.

B. *This Work*

The aim of this work is to find answers to the following questions:

- Do the analyzed radio stations differ regarding the speech rate?
- Is there a relationship between preferences of older adults with regards to the radio stations they listen to and the speech rates at which the announcers at these radio stations speak?
- Does artificial slowing down of the speech signal improve intelligibility for the older adults?

Another aim of this work is to present a review of speech processing algorithms which could potentially be helpful with changing the speech rate.

The paper is arranged as follows: first, in Section II we present some background on how older adults perceive speech. Then, in Section III we describe speech processing algorithms that are able to modify speech rates. In Section IV we describe the setup of the experiments we conducted, followed by the results in Section V. Finally, Section VI presents a discussion of the results and our conclusions.

II. Speech Understanding by Older Adults

There is a substantial amount of research concerning the topic of speech understanding and speech perception by older adults. Humes in [4] analyzed potential reasons for decreased speech understanding by older adults by reviewing experiments conducted in his and other laboratories. For Humes speech understanding meant that the speech signal is “accurately perceived whether in a discrimination, identification, recognition, or comprehension paradigm”. He considered three hypotheses:

- Peripheral – speech understanding is affected by changes in audibility associated with peripheral changes in the human ear;
- Central-auditory – older adults suffer from ”other peripheral deficits accompanying cochlear pathology beyond the loss of hearing sensitivity, such as abnormal spectral and temporal resolution”, caused by functional or structural changes in the auditory pathways between the ear and the cortex;
- Cognitive – speech understanding is affected by deficits in information processing performance by the cortex.

Humes concluded that the strongest evidence exists for the first, peripheral hypothesis, i.e., that hearing loss is to be blamed, such as reduced sensitivity to higher frequencies. The author mentions, however, that “there is increasing evidence that speech unnaturally distorted temporally by means other than reverberation, such as time compression or interruption, may be particularly difficult for the elderly and may not be explained by hearing loss alone.”

Several studies concerns older adults’ perception of fast speech. Letowski and Poch [5] analyzed perception of time compressed speech, both by young and elderly listeners. In their experiments, speech was compressed by regularly discarding speech portions. The authors measured speech understanding scores in the function of the length of Discarded Intervals (DI). Using a group of 15 elderly listeners the authors showed that DI lengths up to 55 msec did not affect the understanding of continuous speech by older adults, but understanding dropped to 40% at 95 msec. In contrast, young listeners demonstrated a tolerance for much longer DIs than the older group – their performance started to decrease significantly as late as 135 msec.

Janse in [6] discussed a somewhat similar problem. In her work, she described experiments with older and young listeners exposed to speech which had been time-compressed either by resampling or by using Time-Domain Pitch Synchronous Overlap And Add (TD-PSOLA) algorithm [7], i.e., by much more advanced methods than the one used by Letowski and Poch [5]. During experiments the subjects were asked to react to a pre-defined target word in the speech signal. Janse found that the Response Time (RT), when increasing speech rates (from 1.5 times the normal rate to 2.0), increased much more for older adult listeners (on average from 314 to 432 ms) than it did for younger listeners (from 199 ms to 249 ms).

The author analyzed potential reasons for the decreased perception of fast speech by older adults, such as hearing loss or information processing rate measures. Janse found that there is a significant correlation between hearing acuity and perception of fast speech by older adults. She also found that speech compressed with TD-PSOLA turned out to be much more intelligible than resampled speech, both for young and older listeners.

In contrast, some researches tried to slow down the speech rate. For example, Gordon-Salant and Fitzgibbons, in their work from 1997 [8], examined whether an increased Inter-Word Interval (IWI), i.e., introducing silence between words, improved speech recognition by older adults. They exposed four different listener groups: young adults and elderly adults, either with or without hearing loss, each consisting of 10 people, to 25 High-Predictability (HP) and 25 low-predictability (LP) sentences. They found that elderly listeners' speech understanding performance did not improve with slowing down the speech rate. However, they noticed that a few elderly individuals with hearing loss showed improvement with speech understanding when IWI was the longest.

Another group of studies concerns the perception of speech by older adults in noisy conditions. For example, Hodoshima *et al*. [9] describe difficulties older adults can encounter in reverberant environments. Pichora-Fuller [10] also discusses speech perception in noise. The author notes that as the signal-to-noise conditions become less favorable, the older listeners enter the zone of "effortful listening" much faster than do young listeners. Having considered this phenomenon, it is easy to understand why in his article Herold [2] instructs "audio engineers" not to add any background noise (such as music, street sound) to speech signal while "performers are speaking important lines." However, he gives no evidence of how harmful background noise is to an elderly listener.

III. Prosodic Modifications

If we search for speech processing algorithms that can potentially help older adults to understand speech, we believe we should focus on the so-called prosodic modification algorithms. The term "prosody" refers to such features of speech as: rhythm, stress or intonation.

According to the commonly accepted source-filter model of speech generation [11], the speech signal s(t) can be represented as:

$$s(t) = g(t) * h(t) \tag{2}$$

where g(t) denotes the excitation signal coming from the larynx (the source signal), h(t) stands for the impulse response of the vocal tract (treated as a filter), and * denotes convolution. While h(t) carries information about speakers individuality, g(t) is mostly affected by speaking style.

There is a family of voice conversion algorithms that aim to change speech spectral features so that it sounds as if speech was produced by someone else. One of the examples of such an algorithm is Gaussian dependent filtering [12]. Such algorithms usually modify only the filter parameters (h(t)). In contrast, prosodic modification algorithms aim at modifying only the source signal component – g(t). In general, they can by divided into three groups [13]:

- Time-scale modification – changing the rate of articulation without affecting speech quality;
- Pitch modification – changing the fundamental frequency (F 0), i.e., the frequency of perceived vocal folds vibration, what results in raising/lowering voice pitch and/or modification of intonation;
- Intensity modification – altering speech energy.

Mixed modifications, e.g., combined time-scale and pitch modifications, are also possible.

From the point of view of this study, time-scale modification, e.g., changing speech rates, seems most suitable to modifying speech in order to potentially help older adults with speech understanding. Moving forward, we will only focus on time-scale modification algorithms.

The oldest methods to change speech rates, used both in experiments and in real-life applications, were rather primitive. One was used to slow down speech and consisted in inserting silence intervals between words – this approach was used, e.g., in [8]. Even though the authors claimed that it had some positive impact on a few listeners, although presumably, listening comfort was seriously impacted as natural speech is predominantly continuous. Another drawback of this approach is that, even though the overall length of utterance is undoubtedly lengthened, the actual speech rate of words remains untouched – length values of phonemes and syllables are not affected. This may be one of the reasons why slowing down speech using this method had rather limited impact on intelligibility [8]. Regardless, modifying silence lengths, e.g., between sentences, can be useful in some cases, to, for instance, synchronize audio with video [14].

One of the methods used in the past to speed up speech was based on overlapping words. It was tested, e.g., in one of the studies described earlier [5]. Other older methods to modify speech rate used resampling, i.e., slowing down playback of speech. Janse [6] employed this method to increase speech rates. Unfortunately, such a method introduces artifacts – it modifies not only speech rates, but also pitch and spectrum, so that speech signal sounds unnatural and annoying; it is not surprising that this method proved to be inferior to other tested methods (e.g., TD-PSOLA, described in the next paragraph).

The TD-PSOLA (Time-Domain Pitch Synchronous Overlap and Add) algorithm is a much more advanced method to modify speech in the time domain without significant spectral modifications. Moulines and Charpentier first proposed this method [7], which is based on copying the speech signal into short-time signals, synchronously with the pitch rate, i.e., synchronously with frequency of vocal folds vibrations. After windowing, these signals are then placed in the new time instants, according to the desired new speech rate, and the overlapping segments are summed together. Thus, if speech is supposed to have a lower speaking rate, the short-time signals are placed at longer distances than in the original signal. This method, together with other variants of OLA (Overlap and Add) techniques, was used in numerous applications and studies, including the before mentioned [6] or [15], and it will also be used in the current experiments.

In the literature, one can find information about trials using speech processing techniques to help older adults with speech understanding. Xiao *et al*. [15] used SOLA-FS (Synchronous Overlap and Add - Fixed Synthesis) [16] algorithm to selectively lengthen consonants in Chinese Mandarin. The authors claimed that four out of five tested patients preferred such altered-rhythm speech, although no intelligibility results were presented. In another example, [14] Nakamura et al. proposed a Portable Real Time Speech Rate Conversion System, which was able to slow down speech to a desired level. The authors used a TDHS (Time Domain Harmonic Scaling) algorithm as a method to change the speech rate without significant quality loss. The authors ran intelligibility tests for the output speech with speech rate changed using rate factors: 0.8, 1.0 (original), 1.2 and 1.4. The researchers presented confusion matrices for Japanese phonemes and two announcers: male and female. Unfortunately, the tests were only conducted with young listeners. In addition, the results showed that almost all phonemes were correctly recognized regardless of the speech rate, thus these results cannot be considered meaningful.

IV. Experiment Design

A. Setup of Speech Rate Measurements

In this study, we decided to analyze six radio stations, which according to Trochimczuk [3] turned out to be most popular among older adult radio listeners in Warsaw. The first three stations (PR1, TOKFM and RM) are listened to by more than 63% of older adults, whilst the next three ones (ZET, PR3 and RMFFM) are listened to by only approximately 17% of older adults and by more than 34% of the young and middle-age audience, so in fact they are already preferred by the younger, and not by older adults listeners.

As one of the objectives of this study was to analyze speech rates of radio stations, we recorded news transmitted by the six study radio stations and analyzed the three first sentences of four speakers presenting news on the air. The recordings took place between 24 and 28 October 2014. Similar to other studies (e.g., [6]), we assessed speech rates by measuring the length (in seconds) of an utterance and manually counted syllables within a time unit (a second), to express speech rates in syllables per second (syl/sec).

B. Setup of Intelligibility Tests

We decided to run intelligibility tests using a method based on Semantically Unpredictable Sentences (SUS) [17]. This method is similar to using low-predictability sentences [8]. The idea of using SUS sentences consists of exposing listeners to utterances which:

- Contain frequent words from a given language;
- Are grammatically correct;
- Make no sense.

A typical example of a SUS sentence would be the following English phrase, “Colorless green ideas sleep furiously” [18]. Due to the semantic unpredictability of these sentences, there is no way that a word can be figured out from the context, unless it is correctly perceived. This method was used in the past to test, for example, quality of speech synthesis [19] and assess quality of voice transmission in the Internet telephony [20], [21].

The audio material contained 50 SUS sentences in Polish, taken from Oliver [22], and was created using five typical sentence structures. The sentences were read by a professional speaker at a high rate of speech – on average 6.8 syllables per second. The sentences were then uniformly slowed down using the TD-PSOLA algorithm implemented in the Praat application [23] and using a time compression factor between 1.15 and 1.5, thus resulting in an average output speech rate of 5.2 syl/sec.

Four participants (native Polish, two female and two male, with higher education, mean age 74.5 yrs) took part in the listening tests. They had no diagnosed hearing loss and were able to use handwriting. Each of the participants was exposed to 25 SUS sentences at original, high speech rate, and 25 SUS sentences at a lowered speech rate, in random order. One male and one female participant listened to each SUS sentence (either “fast” or “slow”). The participants were sitting at a desk in a quiet room in a home environment and were writing down what they heard in a questionnaire. The speech was played back using a standard stand-alone PC-class speaker, to resemble audio quality of an average radio. The time for writing was unlimited, however, the participants were allowed to listen to each utterance only once.

V. Experimental Results

A. Results of Speech Rate Measurements

Table II presents the results of speech rate measurements for the six radio stations in Warsaw, Poland, most popular among the older listeners according to [3]. It shows that by far, RM yielded the lowest speech rate – on average less than five syllables per second, the next station, PR1, was almost one syllable per second faster. PR3, TOKFM and RMFFM turned out to be much faster – more than six syllables per second, with ZET being the fastest: on average almost 6.5 syl/sec.

TABLE II: RESULTS OF SPEECH RATE MEASUREMENTS

Radio station	Speech rate [syl/sec]	Variance
PR1	5.88	0.177
PR3	6.04	0.130
RM	4.92	0.088
RMFFM	6.07	0.075
TOKFM	6.05	0.301
ZET	6.43	0.090

TOKFM and PR1 yielded the highest variance, meaning that those stations employ a variety of speakers: both much “slower” and much “faster” than the average. This also means that a higher population of speakers needs to be investigated in the future to get more precise statistics. In contrast, measurements for RMFFM, ZET and RM exhibit a low variance (less or equal to 0.90), thus meaning that the observed speakers were rather uniformly fast (for RMFFM and ZET) or slow (for RM).

Fig. 1 displays the results of speech rate measurements against the ratio between older and younger audiences, for all six analyzed radio stations. The radio stations shown in the left half of the figure have more young (and middle-age) listeners than older adult listeners (> 65 yrs), while the stations in the right half (TOKFM, PR1, RM) are preferred by older adults. Interestingly, the last two radio stations also exhibit the lowest speech rates. What is more, TOKFM and PR1 yielded rather high speech rate variance, thus meaning that some speakers speak more slowly there as well.

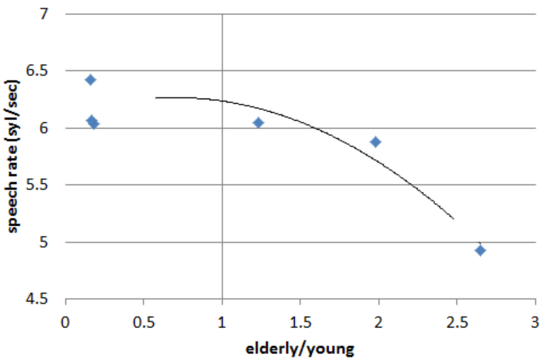

Fig. 1. Speech rate against older/younger audience ratio for the six radio stations analyzed. The solid line represents a hypothetical trend line.

In contrast, stations on the left side, i.e., preferred by the younger audience, yielded the highest speech rates – six or more syllables per second. The line in the figure shows a hypothetical trend between speech rates and age structure of the audience of a given radio station. However, caution must be taken, as only preliminary speech rate measurements were carried out, and only six radio stations were analyzed in this study.

B. *Results of Intelligibility Tests*

We analyzed the sentences written down by participants for the intelligibility tests, considering the following metrics:

- Sentence Recognition Rate (SRR) – percentage of sentences correctly received (i.e., correctly noted by a listener);
- Word Recognition Rate (WRR) – percentage of words correctly received;
- Number of words received in errors.

Naturally, the above metrics are mutually dependent – e.g., a sentence correctly received is a sentence where there are no missed, substituted or inserted words. The results of the intelligibility tests are presented in Table III.

TABLE III: RESULTS OF INTELLIGIBILITY TESTS (SRR – AVERAGE SENTENCE RECOGNITION RATE, WRR – AVERAGE WORD RECOGNITION RATE).

Speech type	WRR [%]	SRR [%]	Word errors
Normal speech	70.0	92.0	39
Slow speech	78.0	94.0	30

The results showed that speech artificially slowed down seems to be more intelligible in terms of the percentage of sentences correctly received (on average, SRR of 78% vs. 70% for original speech), also the WRR was higher when the speech was slow, although the difference here is less distinct. The number of words wrongly received decreased by more than 20% if listeners were exposed to slower speech.

We also analyzed differences between speakers. We found that one of them (the oldest) yielded much higher intelligibility for slow speech than for fast speech (68% vs. 36% SRR), although in general these results were the lowest. The other three participants yielded higher results (around 80% SRR), and showed either a slight preference towards the slower speech or had equal intelligibility for both speech rates. This may suggest that lower speech rates can be potentially more beneficial to listeners with higher hearing deficiencies.

VI. DISCUSSION AND CONCLUSIONS

This paper reviewed the problem of speech intelligibility by the older people. We presented the results of several studies that concerned the problem of speech perception of older adults. As we showed, the results of these studies are not fully consistent – some of the studies (e.g., Gordon-Salant and Fitzgibbons [8]) claim that slowing down speech does not improve intelligibility among older adults, while some other studies develop algorithms to slow down speech or at least lengthen selected phonemes to increase speech understanding by older people (e.g., Xiao *et al.* [15]). Therefore, following the review of the problem of speech understanding by older adults, we presented several speech processing algorithms, which were or potentially could be used in assistive technologies to support older adults.

By using an example of six Polish radio stations, we showed that there is a significant difference for speech rates among various radio stations. We also found that audience ratings for different age groups are, to some extent, correlated with the speech rates of these stations.

We also ran preliminary experiments of intelligibility tests using Semantically Unpredictable Sentences (SUS), exposing the listeners to the original and artificially slowed down speech. While the results seem to be moderately in favor of slower speech, caution must be taken because of the limited number of participants in this test (four) and, therefore, a high margin of error. It is remarkable, however, that a listener who yielded the lowest intelligibility results (90% WRR), benefited most from lowering the speech rate.

We plan to further investigate the issues described in this paper, considering the importance of the topic for our aging society. Potential future works could involve non-uniform lengthening of the speech signal and more profound intelligibility testing.

REFERENCES

[1] Department of Economic and Social Affairs, World Population Ageing 2013. United Nations, December 2013.

[2] E. W. Herold, "Audio for the elderly," *J. Audio Eng. Soc.*, vol. 36, no. 10, pp. 798–800, 1988.

[3] M. Trochimczuk, “Rynek programów radiowych w warszawie–preferencje słuchaczy,” Tech. Rep., National Broadcasting Council, Warsaw, Poland, 2013.

[4] L. E. Humes, “Speech understanding in the elderly,” *Journal of the American Academy of Audiology*, vol. 7, pp. 161–167, 1996.

[5] T. Letowski and N. Poch, “Understanding of time-compressed speech by older adults: Effect of discard interval,” *Journal of the American Academy of Audiology*, vol. 6, no. 6, pp. 433–439, 1995.

[6] E. Janse, “Processing of fast speech by elderly listeners,” *Journal of the Acoustical Society of America*, vol. 125, no. 4, pp. 2361–2373, 2009.

[7] E. Moulines and F. Charpentier, “Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones,” *Speech Communication*, vol. 9, no. 56, pp. 453–467, 1990.

[8] S. Gordon-Salant and P. J. Fitzgibbons, “Selected cognitive factors and speech recognition performance among young and elderly listeners,” *Journal of Speech, Language, and Hearing Research*, vol. 40, no. 2, pp. 423–431, 1997.

[9] N. Hodoshima, T. Arai, and K. Kurisu, “Intelligibility of speech spoken in noise/reverberation for older adults in reverberant environments,” in *Proc. INTERSPEECH*, ISCA, 2012.

[10] M. K. Pichora-Fuller, “Processing speed and timing in aging adults: Psychoacoustics, speech perception, and comprehension,” *International Journal of Audiology*, vol. 42, pp. S59–S67, 2003.

[11] G. Fant, *Acoustic Theory of Speech Production*, Walter de Gruyter, 1970.

[12] J. F. Bonastre, D. Matrouf, and C. Fredouille, “Transfer function-based voice transformation for speaker recognition,” in *Proc. Speaker and Language Recognition Workshop*, 2006, pp. 1–6.

[13] Y. Stylianou, “Voice transformation: A survey,” in *Proc. IEEE International Conference on in Acoustics, Speech and Signal Processing*, 2009, pp. 3585–3588.

[14] A. Nakamura, N. Seiyama, A. Imai, T. Takagi, and E. Miyasaka, “A new approach to compensate degeneration of speech intelligibility for elderly listeners-development of a portable real time speech rate conversion system,” *IEEE Transactions on Broadcasting*, vol. 42, no. 3, pp. 285–293, 1996.

[15] X. Xiao, G. Hu, J. Liu, and C. Liu, “Improving speech intelligibility of older hearing impaired adults by promoting consonant-vowel ratio,” in *Proc. 2nd International Conference on Biomedical Engineering and Informatics*, Oct. 2009, pp. 1–5.

[16] D. J. Hejna Jr, B. R. Musicus, and A. S. Crowe, “Method for time-scale modification of signals,” US Patent 5,175,769, Dec. 29, 1992.

[17] C. Benoit, M. Grice, and V. Hazan, “The SUS test: A method for the assessment of text-to-speech synthesis intelligibility using semantically unpredictable sentences,” *Speech Communication*, vol. 18, no. 4, pp. 381–392, 1996.

[18] N. Chomsky, *Syntactic Structures*, The Hague: Mouton & Co., 1957. Reprinted 1985 by Springer, Berlin and New York.

[19] S. King, “Measuring a decade of progress in text-to-speech,” *Loquens*, vol. 1, no. 1, pp. 6, 2014.

[20] A. Janicki, B. Księżak, S. Kula, and J. Kijewski, “Badanie jakości sygnału mowy w telefonii internetowej z wykorzystaniem zdań nieprzewidywalnych semantycznie,” *Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne*, pp. 257–260, 2006.

[21] A. Raake, *Speech Quality of VoIP: Assessment and Prediction*, John Wiley & Sons, 2007.

[22] D. Oliver, “Diphone synthesis of the Polish language,” M.Sc. thesis, Department of Linguistics, University of Edinburgh, 1998.

[23] P. Boersma, “Praat, a system for doing phonetics by computer,” *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2001.

Artur Janicki received MSc and PhD (1997 and 2004, respectively, both with honors) in telecommunications from the Faculty of Electronics and Information Technology, Warsaw University of Technology (WUT). Assistant Professor at the Institute of Telecommunications, WUT. His research and teaching activities focus on speech processing, including speaker recognition, speech coding and synthesis, emotion recognition, with elements of data mining and information theory. Author of over 30 conference and journal papers, supervisor of over 40 bachelor and master theses. Member of the International Speech Communication Association (ISCA) and the European Association for Signal Processing (EURASIP).

Krzysztof Szczypiorski is Professor of Telecommunications at Warsaw University of Technology, Poland. He holds DSc (habilitation, 2012), PhD (doctorate, 2007) and MSc (master’s degree, 1997) all in Telecommunications from WUT. He also finished his postgraduate studies in Psychology of Motivation (2013) at the University of Social Sciences and Humanities (SWPS), Warsaw, Poland. He graduated from Hass School of Business, University of California, Berkeley, USA (2013). Head and founder of: Cybersecurity Division at Institute of Telecommunications, WUT. A research leader of Network Security Group at WUT. His research interests include: theory of observing change, network security, digital forensics, open-source intelligence, and wireless and ad-hoc communications. He is the author or the co-author of 150+ papers and 50+ invited talks. He is the inventor of 2 patents (one of them is pending).

The guest editor of 10+ special issues of top ICT journals. For 20+ years he serves as the independent consultant in fields of network security, telecommunications and computer science for many entities including: Cisco Systems, Hewlett-Packard, Ministry of Finance (Poland), Oracle, Orange Poland, Parliament of Republic of Poland, Polish Energy Group, PwC, T-Mobile Poland. Personal website: http://ksz.tele.pw.edu.pl.