ARCHIVES OF ACOUSTICS
Vol. **44**, No. 3, pp. 439–446 (2019)
DOI: 10.24425/aoa.2019.129257

Acoustic Parameters in the Evaluation of Voice Quality of Choral Singers. Prototype of Mobile Application for Voice Quality Evaluation

Krzysztof SZKLANNY

Multimedia Department
Polish-Japanese Academy of Information Technology
Koszykowa 86, 02-008 Warsaw, Poland; e-mail: kszklanny@pjwstk.edu.pl

(*received January 30, 2019; accepted May 6, 2019*)

Choral singers are among intensive voice users whose excessive vocal effort puts them at risk of developing voice disorders. The aim of the work was to assess voice quality for choral singers in the choir at the Polish-Japanese Academy of Information Technology. This evaluation was carried out using the acoustic parameters from the COVAREP (A Collaborative Voice Analysis Repository For Speech Technologies) repository. A prototype of a mobile application was also prepared to allow the calculation of these parameters.

The study group comprised 6 male and 19 female choir singers. The control group consisted of healthy non-singing individuals, 50 men and 39 women. Auditory perceptual assessment (using the RBH scale) as well as acoustic analysis were used to test the voice quality of all the participants.

The voice quality of the female choir singers proved to be normal in comparison with the control group. The male choir singers were found to have tense voice in comparison with the controls. The parameters which proved most effective for voice evaluation were Peak Slope and Normalized Amplitude Quotient.

Keywords: web application; voice analysis; voice quality; acoustic analysis; COVAREP.

1. Introduction

According to the UEP guidelines (the Union of the European Phoniatricians), occupations with high vocal load can be divided into three groups. Group 1 includes professions that require a special voice quality, such as vocal performers, choral singers, actors, radio and television presenters. Group 2 comprises occupations that involve a considerable strain on people's vocal apparatus, such as educators, interpreters, politicians, call center workers, customer service assistants. Group 3 consists of occupations that require higher-than-average vocal demands, such as lawyers and army officers (Pruszewicz, 1992; Niebudek-Bogusz, Śliwińska-Kowalska, 2006). Substantial vocal effort is often caused by inappropriate working conditions or various stressors. These often have a considerable effect on people's physical condition, including voice quality. The most important risk factors that can lead to the development of occupational voice disorders include incorrect vocal habits, voice overuse, negative personality traits (psychosomatic factors) such as high-conflict personality and nervousness, prolonged exposure to stress, poor working conditions and failure to observe the basic rules of vocal hygiene (e.g. smoking) (Gundermann, 1970; NIDOCD, 2016).

According to the protocol devised by the European Laryngological Society (ELS), perceptual and acoustic analyses of voice quality are the basic tools for diagnosing voice disorders (Dejonckere *et al.*, 2003). Acoustic analysis has the advantage of being objective and noninvasive. By applying a parametrized acoustic signal one obtains information about voice quality (Marasek, 1997). There are a number of parameters available for the evaluation of voice quality. One of the best implementations of acoustic parameters is COVAREP (A Collaborative Voice Analysis Repository For Speech Technologies) (Degottex *et al.*, 2014).

The research objective was to establish whether the voice quality of choral singers differs from the voice quality of healthy non-professional voice users by applying perceptual and acoustic analysis. Types of voice-quality abnormalities were also examined, on the basis of recordings of singers from the Polish-Japanese

Academy of Information Technology choir. The aim of this study was to develop a prototype web app for physicians, choral singers and speech researchers which would allow them to check voice quality while using algorithms implemented in the COVAREP repository.

2. Web and mobile application for the analysis of voice quality

There are a number of mobile applications that are useful for diagnosing communication disorders (Stasak, Epps, 2017; Ventola, 2014; Gravenhorst *et al.*, 2015; Miloff *et al.*, 2015; Beiwinkel *et al.*, 2016; Nicholas *et al.*, 2015). However, only some of them can be used to evaluate voice quality. Research study (Verde *et al.*, 2015) presents an app that enables the measurement of F0 – the fundamental frequency of the vocal folds. The app analyzes a sustained phonation of the vowel /a:/ and produces a report with reference to voice quality. At present, the app is no longer available.

Another app of this kind is OperaVOX. It provides highly consistent measurements of F0, Jitter, Shimmer and NHR (Noise-to-Harmonics Ratio) (Mat Baki *et al.*, 2015). Likewise, this particular app is no longer available.

Another app for the analysis of voice quality is presented in a study by van Leer *et al.* (2017). It works in iOS and makes it possible to calculate the CPP parameter (Cepstral Peak Prominence). It is dedicated to voice therapy patients due to their persistently incorrect vocal habits. The app allows for the calculation of a parameter that enables the achievement of resonant voice production.

People wanting to alter the pitch or volume of their voice, including those with dysphonia or Parkinson's disease, can use the Voice Analyst app, which will provide them with useful visual, acoustic and statistical feedback about their voice. The app offers the possibility of saving the recorded material and can also be used by singers who want to improve the acoustic quality of their voice (Voice Analyst app, 2019).

Voice Online Lab, an app developed by Voice Clinical Systems, is available in Google Play. It is designed to automate and facilitate the collection of patients' voice samples with the help of mobile devices. Unfortunately, Voice Online Lab can only be used in three countries – the UK, France and Spain. Programmed on the basis of the external-client – central-server model, the app functions in accordance with the following algorithm: a physician registers on the Voice Clinical Systems portal, where he/she can create temporary accounts for his/her patients and analyze the results of their voice recordings; the patient downloads the app in the Android system and logs on to the previously generated account; at the app level, the patient records his/her voice sample and completes a short form providing information about his/her sex, age and current state of health; the saved file is then analyzed in terms of noise as well as the quality and length of the recording; if the recording meets the required standard, the patient selects one of the three types of report – a simple abridged report, a simple detailed report or a comprehensive report; once the type of report is selected, a relevant fee in Euros is charged; when the right amount has been paid, the sample is sent over to the central Voice Clinical Systems server, where it is analyzed and where the required parameters, previously defined by the doctor, are calculated; following the analysis, the sample is sent to the patient's doctor who interprets, and possibly corrects, the results; the approved and interpreted results, along with possible comments from the doctor, are then sent to the patient. The app makes use of the acoustic parameters implemented in the COVAREP repository (Voice Online Lab, 2019).

3. Selection of acoustic parameters for web app implementation

COVAREP (A Collaborative Voice Analysis Repository For Speech Technologies) is currently one of the best implementations of acoustic parameters of the human voice (Degottex *et al.*, 2014). The following acoustic parameters are implemented in the repository: Peak Slope (PS) (Kane, Gobl, 2011), H1H2 – the amplitude difference between the first and the second harmonics (Titze, Sundberg, 1992), Normalized Amplitude Quotient (NAQ) (Alku *et al.*, 2002), Parabolic Spectral Parameter (PSP) (Alku *et al.*, 1997), Quasi Open Quotient – QOQ (Hacki, 1989), Cepstral Peak Prominence (CPP) (Hillenbrand *et al.*, 1994), Harmonic Richness Factor (HRF) (Childers, Lee, 1991). These parameters were all implemented in the web app WEBAA presented in this study. They were then calculated for both the choir singers and the control group. Below are the definitions of the parameters that are most important from the perspective of this study.

3.1. Peak Slope

Peak Slope (PS), a parameter defined through a wavelet transform, helps to distinguish between modal voice and tense or breathy voice. The following definition of the mother wavelet function is used:

$$g(t) = -\cos(2\pi f_n t)\cdot \exp\left(\frac{-t^2}{2\tau^2}\right), \tag{1}$$

where $f_s = 16$ kHz, $f_n = \frac{f_s}{2}$ and $\tau = \frac{1}{2f_n}$.

Formula (1) shows the mother wavelet used to calculate the PS parameter value. The analysed speech signal recorded with an $x(t)$ microphone is decomposed through the use of convolution $g\left(\frac{t}{s_i}\right)$, where $s_i = 2^i$, $i = 0, 1, 2, ..., 5$. The result is a division into octave bands with a centre frequency of 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz. Subsequently, a local maximum for each audio signal filter is calculated. The next step involves calculating the linear regression for the maxima. Peak Slope is robust to babble noise with a signal-to-noise ratio of 10 dB, which is critical in speech signal recordings. The results obtained in a study by Kane confirm the parameter's reliability in distinguishing between modal voice and tense or breathy voice (Kane, Gobl, 2011).

3.2. Normalized Amplitude Quotient

Normalized Amplitude Quotient (NAQ) is a time-based parameter of speech signal analysis. It is computed for each period of glottal flow using the following formula (2) (Alku *et al.*, 2002):

$$\frac{A_{ac}}{T_{av} d_{\min}} = \frac{A_{\max} - A_{\min}}{T_{av} d_{\min}}, \tag{2}$$

where $A_{\max}$ is amplitude for each period of the signal, $A_{\min}$ is the lowest amplitude for each period of the signal, T_{av} – the average fundamental period length, $d_{\min}$ minimum derivative glottal flow, and A_{ac} – maximal flow of amplitude. The efficacy of NAQ in differentiating between phonation types was validated in prior works by Bäckström, Alku and Kane (Bäckström *et al.*, 2002; Kane, Gobl, 2011).

4. Implementation of the web app

WEBBA, the app developed in this study, is a web app with a GUI designed for mobile devices. Thanks to it, one can load an audio file recorded with a smartphone, select acoustic parameters for calculations and save them in a JSON, CSV, PNG or PDF file. The users include the administrator, doctors and patients. Data concerning patients' medical condition and enabling the identification of a patient are protected with particular care. They are retained pursuant to European Parliament and European Council Regulation (EU) 2016/679 of April 27, 2016. The authors must also comply with the Polish data protection act (*Personally Identifiable Information Protection Act*, Dz.U. 2018 poz. 1000 of May 10, 2018).

4.1. System architecture

The WEBAA app is divided into components and modules (Wilson, 2016). This kind of architecture makes it easier to optimize and implement the functionality of the system.

The first element of the architecture is the access layer, also known as the external layer. It consists of all the hardware/software tools that allow the database app to be run and accessed by an external user.

The access layer consists of:

- a WWW server (NGiNX),
- a cache server (Redis),
- a Fail2ban application,
- a FirewallD application.

The WWW server is responsible for delivering the presentation module to the user. The Cache server is responsible for relieving the main database of repeatedly recurring requests. It also stores the sessions of logged-in users. Fail2ban protects the server from troublesome brute-force attacks on SSH. FirewallD handles access to TCP/IP ports.

The presentation module handles clients' requests and presents the content in a manner accessible to the end user. The module was written in PHP and the Laravel 5.5 LTS framework. The presentation module and the submodules were implemented using the Action-Domain Responder pattern (ADR) and Hexagonal Architecture (Buenosvinos *et al.*, 2017; Corona, 2014). ADR is a refined version of the commonly known and used Model-View-Controller pattern (Krasner, Pope, 1988). The presentation component is responsible for delivering the contents requested by a client. In most cases, these will be webpages (HTML) and webservices (SOAP), REST (JSON).

The authorization module is responsible for the authentication of system users. A user will not be able to process files without having a valid account. In this app, the authorization is based on a Laravel component that handles session data and the login.

The module is responsible for the handling of patient data (creating, browsing, searching, deleting), e.g. adding, deleting or editing information about a patient, as well as providing information about the time of the recording being processed.

The role of the module is to process and analyze .wav files. The calculation module operates using Audio Toolkit – a set of algorithms and classes designed to facilitate the handling and resampling of .wav files, and Signal Toolkit – a library that enables signal processing through noise reduction, signal behavior prediction and signal filtering.

4.2. Tests

To validate the correct performance of the software and to check whether the predefined specifications were met, a number of unit and acceptance tests were carried out. Acceptance tests can be written either as a code or as test scenarios which are then executed by independent testers. The unit tests were written using the specialized PHP Unit library (PHP, 2019). The WEBAA app was verified with both unit tests and scenario-based acceptance tests. Subsequently, WEBAA was installed and configured in a test environment. The test scenarios were executed step by step by two testers. Tester 1 tested sections 1.0–3.0, whereas Tester 2 tested sections 3.1–5.0. The tests were conducted in Mozilla Firefox 61.01 and Google Chrome 68. The following functionalities of the system were tested:

- 1.0 – System login,
- 1.1 – System logout,
- 2.0 – Home Page user logged in,
- 2.1 – Access to system subpages without login,
- 2.2 – Display of non-existing subpage,
- 3.0 – Display of examination,
- 3.1 – Adding new examination,
- 4.0 – Adding new sample,
- 4.1 – Display of sample results,
- 4.2 – Downloading sample results,
- 5.0 – Adding new patient.

The app is available at https://webaa.thebb.pl.

5. Recordings of academic choir and control group

The next stage involved the verification of the app's effectiveness in assessing choral singers' voice quality. The academic choir recordings were made in the recording studio of the Polish-Japanese Academy of Information Technology. The group consisted of 6 male and 19 female choir singers whose average age was 34.64 and 36.08, respectively. In the control group it was 35.57 and 31.8. The choir's voice classification was as follows: 3 basses, 2 baritones, 1 tenor, 10 sopranos and 9 altos. The control group was made up of 50 males and 39 females.

The microphone used in the recordings was a Rode NT-1A and it has following parameters; frequency range 20 Hz – 20 kHz, sensitivity 25 mV/Pa; equivalent Noise Level 5 dBA, maximum SPL – 137 dB SPL, polar pattern – cardioid.

The signal was registered with a 48 kHz sampling rate and a 16-bit resolution (standard WAV PCM). During the recording, the choir singers and the controls phonated the vowel /a:/ three times with a sound pressure level of 60–80 dBA, 1m from the microphone, for a sustained period of at least 4 seconds. Following that, the recorded individual was made to briefly strain his/her voice by reading out a few sentences, and then again to phonate the vowel /a:/ four times. The last four phonations of the vowel /a:/ were used to calculate acoustic parameters. All the participants phonated in a neutral manner. Phonations with higher or lower values of F0 were not taken into account in the analyses.

The microphone used in this research has better specification than the one built into the smartphone. However, a test was carried out involving the recording of choral singer using Sony Xperia XZ Premium smartphone and an external PreSonus AudioBox USB 96 interface connected to it. A Rode NT-1A microphone was attached to the audio interface. At the same time, another recording was made using the Rode NT-1A microphone connected to the computer. In the 20 recordings of the vowel /a:/, there were no differences in COVAREP parameters values (Sec. 3).

6. Perceptual assessment of voice quality

The examinations were supplemented by voice quality evaluation based on a perceptual assessment of voice quality on the RBH scale (Nawka *et al.*, 1994). The sentences had been constructed in a way that would facilitate such assessment. The RBH perceptual scale is used in German clinics and is recommended by the Committee on Phoniatrics of the European Laryngological Society (Dejonckere *et al.*, 2001). The RBH acronym stands for the following features:

R – Rauigkeit (roughness) – the degree of voice roughness caused by irregular vocal fold vibrations,

B – Behauchtheit (breathiness) – the degree of breathiness caused by glottic insufficiency,

H – Heiserkeit (hoarseness) – the degree of hoarseness.

Scores of 0, 1, 2, and 3 are used for all the parameters on the RBH scale, with reference to the different degrees of vocal disorder: '0' = normal voice, '1' = a slight degree, '2' = a medium degree, and '3' = a high degree. Perceptual assessment of voice quality was carried out on both occasions by the same two independent voice specialists who had completed an RBH training program and had extensive experience in voice/speech signal assessment. On both occasions, the experts were blinded for the duration of the assessment.

7. Perceptual assessment results

The statistical analysis was carried out in the Octave environment. The comparison of the accuracy and reliability of the perceptual voice quality assessment on the RBH scale was made by two voice experts using the Mann-Whitney nonparametric test, because the distribution for the RBH scale is not normal. In the experts' annotations, the differences for the parameters R, B and H were not statistically significant (Chi-squared test, p-value >0.05). The experts had no hearing impairment, which was confirmed by an audiometric test.

The p-value for the female choir singers and the control group is 0.8634 for R, 1.0 for B and 0.8724 for H. The p-value for the male choir singers and the control group is 0.6366 for R, 1.0 for B and 0.8997 for H. That being so, the results of the perceptual statistical analysis do not indicate differences between the respective groups. Table 1 shows the distribution of the features indicating changes in voice quality on the RBH scale for the male and female choir singers and the corresponding control groups. In all other cases the feature values were “0”.

Table 1. Voice abnormalities in the male and female choir groups and the corresponding control groups. Male choir singers = 6, female choir singers = 19, male control group = 50, female control group = 39.

Feature distribution	R1	B1	H1
In the male choir group	3	–	1
In the female choir group	1	–	2
In the male control group	2	–	2
In the female control group	1	–	3

8. Acoustic analysis results

Anderson-Darling and Saphiro-Wilk tests were used to verify the normal distribution for each acoustic parameter. All the values of the acoustic parameters have a normal distribution. An F-test was run to check whether the variants were equal.

The differences between the fundamental frequency for the male and female choir singers and the corresponding control groups are not statistically significant (Table 2). As the distribution is not normal, a Mann-Whitney test was run for verification purposes. Calculations for each individual were made on the basis of three phonations. For the male singers and the corresponding control group the p-value = 0.93, U = 411.5, while for the female choir singers and the corresponding control group the p-value = 0.44, U = 945. The differences in values of acoustic parameters were verified with a Student's t-test. Tables 3 and 4 show results of the performed tests.

Table 2. Results of Mann-Whitney test for F0 values.

	Mean	Median	STD	Mann-Whitney test
F0 for male choir singers	111.92	105	17.94	p-value = 0.93, U = 411.5
F0 for male control group	108.32	108	10.47	
F0 for female choir singers	214.02	205	23.28	p-value = 0.44, U = 945
F0 for female control group	209.22	206	16.85	

Table 3. Results of Student's t-test for male choir singers and control group.

	PS	NAQ	CPP	H1H2	PSP	QOQ	HRF
Mean ± SD for male choir singers	−0.389 ± 0.046	0.140 ± 0.037	11.170 ± 0.305	4.486 ± 2.975	0.207 ± 0.071	0.492 ± 0.087	29.227 ± 5.242
Mean ± SD for male control group	−0.332 ± 0.034	0.102 ± 0.021	11.102 ± 0.410	2.152 ± 1.816	0.133 ± 0.036	0.411 ± 0.088	30.489 ± 7.498
t	6.078	5.459	0.803	4.157	5.690	4.006	0.834
t, critical	1.993	1.995	1.993	1.993	1.995	1.993	1.993
p-value	<0.0001	<0.0001	0.4244	<0.0001	<0.0001	0.0001	0.4072

Table 4. Results of Student's t-test for female choir singers and control group.

	PS	NAQ	CPP	H1H2	PSP	QOQ	HRF
Mean ± SD for female choir singers	−0.281 ± 0.084	0.121 ± 0.031	11.729 ± 0.245	9.125 ± 3.573	0.210 ± 0.044	0.417 ± 0.069	15.954 ± 2.804
Mean ± SD for female control group	−0.301 ± 0.041	0.124 ± 0.024	11.786 ± 0.333	8.445 ± 3.680	0.195 ± 0.046	0.394 ± 0.051	17.360 ± 1.750
t	1.496	0.502	1.066	1.012	1.645	1.994	3.098
t, critical	1.983	1.985	1.981	1.980	1.985	1.980	1.980
p-value	0.1377	0.6170	0.2889	0.3137	0.1033	0.0485	0.0024

9. Discussion

The acoustic analysis results confirm that with the help of acoustic analysis it is possible to distinguish between the voice of choral singers and the voice of healthy non-professional voice users. This would not be possible using a perceptual scale, despite being able to indicate changes in voice quality in particular individuals. For example, roughness was observed in three male choir singers and one female choir singer, and hoarseness was observed in one male choir singer and two female choir singers. However, unlike acoustic analysis results, perceptual analysis results do not lead to further, more specific analysis.

The acoustic analysis results indicate that it is possible to distinguish between the voice of female choir singers and the corresponding control group using QOQ (p-value = 0.0485) and HRF (p-value = 0.0024). For the male choir singers and the corresponding control group, the parameters that allow for the voice differentiation are PS (p-value < 0.0001), NAQ (p-value < 0.0001), H1H2 (p-value < 0.0001), PSP (p-value < 0.0001), and QOQ (p-value < 0.0001).

Of the parameters mentioned above, PS and NAQ prove to be the most effective (Kane, Gobl, 2011; Alku *et al.*, 2002). PS is characterized by the highest noise resistance while displaying greater voice quality identification accuracy than the other parameters. PS is completely standalone, i.e. no other algorithms (e.g., F0, GCI detection, inverse filtering) are needed in order to obtain the required values, which is its considerable advantage. Besides, Peak Slope was developed without assumptions which would affect the decision on windowing, which can be quite complicated particularly when analyzing non-modal voices. The effectiveness of PS at 25 dB SNR in differentiating between the three types of phonation (breathy voice, modal voice and tense voice) is approximately 77%. By comparison, the value for NAQ is approximately 68% (Kane, Gobl, 2011).

In the male choir singers group, the PS parameter made it possible to detect a shift in voice quality towards tense voice in comparison with the control group (PS = −0.39 *vs* −0.33). The lower the value of Peak Slope is, the voice becomes tense (Kane, Gobl, 2011; 2013). This may indicate an incorrect vocal technique and, as a result, can lead to voice problems. Choral singers are far more at risk of developing voice disorders than those who do not use their voice for working. Their extreme vocal loads result in overstraining the voice mechanism. Excessive tension of the muscles in the vocal apparatus causes discomfort or pain when speaking. Monitoring shifts in phonation may help choral singers to modify their incorrect vocal habits. Professional voice training courses offer an opportunity to acquire correct vocal habits through intense routine of voice-related muscle conditioning and endurance practice (Vaiano *et al.*, 2013; Braun-Janzen, Zeine, 2009). Owing to that, one can minimize the risk of voice damage. Also, good vocal hygiene will help to maintain the correct functioning of the vocal apparatus. A study by Siupsinskiene and Lycke (2011) confirmed that voice training has a significantly positive effect on vocal capabilities, mostly singing voice.

NAQ, too, was found to be an effective parameter in differentiating between different types of phonation (Kane, Gobl, 2011; 2013). Study (Björkner *et al.*, 2006), however, concluded that in the case of singers this particular parameter may yield less reliable results. It was found that high NAQ values indicate a less abducted phonation type, while lower values indicate a more abducted phonation type. It is also likely that well-trained singers try to retain the characteristics of the voice source unaffected by F0 and loudness variation such that they avoid pressed phonation at high pitch and/or high loudness. However, because NAQ assumes different values between F0 278 Hz and 139 Hz, it is essential that the F0 values for singers and the corresponding controls should not be statistically significant. Test results show that for singers the parameter requires further experiments, which coincides with the author's hypothesis. Choral singers demonstrate considerably better vocal control than their respective controls. This affects the flow of air through the vocal tract and is reflected in the parameter values. PS and NAQ have been effectively implemented in the author's previous studies concerning rare genetic disorders such as Pompe disease and Morquio syndrome type IV, as well as a study devoted to the classification of vocal nodules using genetic algorithm (Szklanny *et al.*, 2016; Szklanny, Tylki-Szymańska, 2018; Szklanny *et al.*, 2018; 2019). In the author's study concerning vocal nodules in children, NAQ accurately reflected the characteristics of tense voice.

The other acoustic parameters either indicate modal voice, or require further research to provide a more precise answer regarding voice quality.

The developed prototype of a web app allows for the calculation of acoustic parameters of voice quality. It is recommended that an external audio interface connected to a smartphone be used to obtain reliable values. In subsequent tests, it is necessary to check how the values of acoustic parameters recorded with a smartphone differ from those parameters recorded with a professional microphone. Such a solution would be more convenient and would not require connecting the audio interface to the smartphone.

10. Conclusion

The conducted studies indicated the existence of abnormal voice qualities in the academic choir of the Polish-Japanese Academy of Information Technology in comparison with the control group. The male singers were found to have tense voice while the female singers showed no significant changes in voice quality. Of the implemented parameters, Peak Slow proved to be the most effective. Not only is it noise-resistant, it also allows for reliable evaluation of voice quality. WEBBA, the web app developed in this study, allows, inter alia, for the measurement of this parameter.

Acknowledgments

I wish to thank Mr. Mateusz Smolik for his assistance in the implementation of the app in the Octave environment.

References

1. Alku P. (2011), *Glottal inverse filtering analysis of human voice production – A review of estimation and parameterization methods of the glottal excitation and their applications*, Sadhana – Academy Proceedings in Engineering Sciences, **36**, 5, 623–650, https://doi.org/10.1007/s12046-011-0041-5.
2. Alku P., Bäckström T., Vilkman E. (2002), *Normalized amplitude quotient for parametrization of the glottal flow*, The Journal of the Acoustical Society of America, **112**, 2, 701–710, https://doi.org/10.1121/1.1490365.
3. Alku P., Strik H., Vilkman E. (1997), *Parabolic spectral parameter – A new method for quantification of the glottal flow*, Speech Communication, **22**, 1, 67–79, https://doi.org/10.1016/S0167-6393(97)00020-4.
4. Bäckström T., Alku P., Vilkman E. (2002), *Time-domain parameterization of the closing phase of glottal airflow waveform from voices over a large intensity range*, IEEE Transactions on Speech and Audio Processing, **10**, 3, 186–192, https://doi.org/10.1109/TSA.2002.1001983.
5. Behlau M., Oliveira G. (2009), *Vocal hygiene for the voice professional*, Current Opinion in Otolaryngology and Head and Neck Surgery, https://doi.org/10.1097/MOO.0b013e32832af105.
6. Beiwinkel T. *et al.* (2016), *Using smartphones to monitor bipolar disorder symptoms: a pilot study*, JMIR Mental Health, **3**, 1, e2, https://doi.org/10.2196/mental.4560.
7. Björkner E., Sundberg J., Alku P. (2006), *Subglottal pressure and normalized amplitude quotient variation in classically trained baritone singers*, Logopedics Phoniatrics Vocology, **31**, 4, 157–165, https://doi.org/10.1080/14015430600576055.
8. Braun-Janzen C., Zeine L. (2009), *Singers' interest and knowledge levels of vocal function and dysfunction: Survey findings*, Journal of Voice, **23**, 4, 470–483, https://doi.org/10.1016/j.jvoice.2008.01.001.
9. Buenosvinos C., Soronellas Ch., Akbary K. (2017), *Domain-Driven Design in PHP*, Packt Publishing.
10. Childers D.G., Lee C.K. (1991), *Vocal quality factors: Analysis, synthesis, and perception*, The Journal of the Acoustical Society of America, **90**, 5, 2394–2410, https://doi.org/10.1121/1.402044.
11. Corona S. (2014), *Scaling PHP7 applications*, Leanpub.
12. Degottex G., Kane J., Drugman T., Raitio T., Scherer S. (2014), *COVAREP – A collaborative voice analysis repository for speech technologies*, [in:] Proceedings of ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing, 4–9 May, Florence, Italy, https://doi.org/10.1109/ICASSP.2014.6853739.
13. Dejonckere P.H. *et al.* (2001), *A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques: Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS)*, European Archives of Oto-Rhino-Laryngology, **258**, 2, 77–82, https://doi.org/10.1007/s004050000299.
14. Dejonckere P.H., Crevier-Buchman L., Marie J.P., Moerman M., Remacle M., Woisard V. (2003), *Implementation of the European Laryngological Society (ELS) basic protocol for assessing voice treatment effect*, Revue de laryngologie-otologie-rhinologie, **124**, 5, 279–283.
15. Gravenhorst F. *et al.* (2015), *Mobile phones as medical devices in mental disorder treatment: an overview*, Personal and Ubiquitous Computing, **19**, 2, 335–353, https://doi.org/10.1007/s00779-014-0829-5.
16. Gundermann H. (1970), *The professional dysphonia* [in German: *Die Berufsdysphonie*], Thieme, Leipzig.
17. Hacki T. (1989), *Classification of glottic functions by means of electroglottography* [in German: *Klassifizierung von Glottisfunktionen mit Hilfe der Elektroglottographie*], Folia Phoniatrica, **41**, 43–48, https://doi.org/10.1159/000265931.
18. Hillenbrand J., Cleveland R.A., Erickson R.L. (1994), *Acoustic correlates of breathy vocal quality*, Journal of Speech Language and Hearing Research, **37**, 4, 769–778, https://doi.org/10.1044/jshr.3704.769.
19. Kane J., Gobl C. (2011), *Identifying regions of non-modal phonation using features of the wavelet transform*, [in:] Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, pp. 177–180.
20. Kane J., Gobl C. (2013), *Wavelet maxima dispersion for breathy to tense voice discrimination*, IEEE Transactions on Audio, Speech and Language Processing, **22**, 6, 1170–1179, https://doi.org/10.1109/TASL.2013.2245653.
21. Krasner G.E., Pope S.T. (1988), *A description of the model-view-controller user interface paradigm in the Smalltalk-80 System*, Journal of Object Oriented Programming, **1**, 3, 26–49.

22. Marasek K. (1997), *Electroglottographic Description of voice quality*, Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung, **3**, 2, Diss. Habilitationsschrift, Stuttgart.
23. Mat Baki M. *et al.* (2015), *Reliability of OperaVOX against Multidimensional Voice Program (MDVP)*, Clinical Otolaryngology, **40**, 22–28, https://doi.org/10.1111/coa.12313.
24. Miloff A., Marklund A., Carlbring P. (2015), *The challenger app for social anxiety disorder: New advances in mobile psychological treatment*, Internet Interventions, **2**, 4, 382–391, https://doi.org/10.1016/j.invent.2015.08.001.
25. Nawka T., Anders L.C., Wendler J. (1994), *The auditory assessment of hoarse voices according to the RBH system* [in German: *Die auditive Beurteilung heiserer Stimmen nach dem RBH-System*], Sprache, Stimme, Gehör, **18**, 130–133.
26. Nicholas J., Larsen M.E., Proudfoot J., Christensen H. (2015), *Mobile apps for bipolar disorder: A systematic review of features and content quality*, Journal of Medical Internet Research, **17**, 8, e198, https://doi.org/10.2196/jmir.4581.
27. NIDOCD – The National Institute on Deafness and Other Communication Disorders (2016), *Statistics on voice, speech, and language*, https://www.nidcd.nih.gov/health/statistics/statistics-voice-speech- and-language#1, July.
28. Niebudek-Bogusz E., Sliwińska-Kowalska M. (2006), *Applicability of voice acoustic analysis with vocal loading test to diagnostics of occupational voice diseases* [in Polish: *Ocena przydatności analizy akustycznej z zastosowaniem próby obciążeniowej w diagnostyce chorób zawodowych narządu głosu*], Medycyna Pracy, **57**, 6, 497–506.
29. *Personally Identifiable Information Protection Act* [in Polish: Ustawa z dnia 10 maja 2018 r. o ochronie danych osobowych], Dz.U. 2018 poz. 1000 of May 10, 2018.
30. PHP The Right Way (2019), http://www.phptherightway.com, access in January 2019.
31. Pruszewicz A. (1992), *Professional voice disorders* [in Polish: *Zawodowe zaburzenia głosu*], [in:] *Foniatria kliniczna*, Państwowy Zakład Wydawnictw Lekarskich, Warszawa, pp. 205–209.
32. Siupsinskiene N., Lycke H. (2011), *Effects of vocal training on singing and speaking voice characteristics in vocally healthy adults and children based on choral and nonchoral data*, Journal of Voice, **25**, 4, e177–e189, https://doi.org/10.1016/j.jvoice.2010.03.010.
33. Stasak B., Epps J. (2017), *Differential performance of automatic speech-based depression classification across smartphones*, [in:] 2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), pp. 171–175, 23–26 October, San Antonio, TX, USA, https://doi.org/10.1109/ACIIW.2017.8272609.
34. Szklanny K., Gubrynowicz R., Iwanicka-Pronicka K., Tylki-Szymańska A. (2016), *Analysis of voice quality in patients with late-onset Pompe disease*, Orphanet Journal of Rare Diseases, **11**, 99, https://doi.org/10.1186/s13023-016-0480-5.
35. Szklanny K., Gubrynowicz R., Ratyńska J., Chojnacka-Wądołowska D. (2019), *Electroglottographic and acoustic analysis of voice in children with vocal nodules*, International Journal of Pediatric Otorhinolaryngology, **22**, 82–88, https://doi.org/10.1016/j.ijporl.2019.03.030.
36. Szklanny K., Gubrynowicz R., Tylki-Szymańska A. (2018), *Voice alterations in patients with Morquio A syndrome*, Journal of Applied Genetics, **59**, 1, 73–80, https://doi.org/10.1007/s13353-017-0421-6.
37. Szklanny K., Tylki-Szymańska A. (2018), *Follow-up analysis of voice quality in patients with late-onset Pompe disease*, Orphanet Journal of Rare Diseases, **13**, 189, https://doi.org/10.1186/s13023-018-0932-1.
38. Titze I.R., Sundberg J. (1992), *Vocal intensity in speakers and singers*, The Journal of the Acoustical Society of America, **91**, 2936, https://doi.org/10.1121/1.402929.
39. Vaiano T., Guerrieri A.C., Behlau M. (2013), *Body pain in classical choral singers*, CoDAS, **25**, 4, 303–309.
40. van Leer E., Pfister R.C., Zhou X. (2017), *An iOS-based cepstral peak prominence application: feasibility for patient practice of resonant voice*, Journal of Voice, **31**, 1, 131.e9-131.e16, https://doi.org/10.1016/j.jvoice.2015.11.022.
41. Ventola C.L. (2014), *Mobile devices and apps for health care professionals: uses and benefits*, P & T: A Peer-Reviewed Journal for Formulary Management, **39**, 5, 356–364.
42. Verde L., De Pietro G., Veltri P., Sannino G. (2015), *An m-health system for the estimation of voice disorders*, [in:] 2015 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2015, 29 June – 3 July, Turin, Italy, pp. 1–6, https://doi.org/10.1109/ICMEW.2015.7169766.
43. Voice Analyst app, Google Play, https://play.google.com/store/apps/details?id=co.speechtools.voiceanalyst&hl=pl, access in January 2019.
44. Voice Online Lab, Google Play, https://play.google.com/store/apps/details?id=com.voicecs.onlinelab&hl=en_US, access in January 2019.
45. Wilson K. (2016), *The clean architecture in PHP*, Lean Publishing.