DOI: 10.21008/j.0860-6897.2023.2.11
Vibrations in Physical Systems, 2023, 34(2), 2023211

Singing voice analysis on the basis of acoustic parameters

Natalia BEDNARZ[1], **Zdzisław MADEJ**[2]

[1] AGH University of Krakow, Department of Mechanical Engineering and Robotics
[2] The Krzysztof Penderecki Academy of Music in Krakow, Department of Voice and Drama

Corresponding author: Natalia BEDNARZ, email: bednarznatalia54@gmail.com

Abstract Voice plays a fundamental role in human relations. In addition to its communicative function in everyday life, the voice also acts as an instrument or a working tool for singers, teachers or actors. It is said that singing is an extension of speech, but performing it correctly is a complex task that requires hard work and training. This paper draws attention to the problem of insufficient training in voice emission and voice control of singers in amateur choirs, which can cause strain and disorders of the phonatory system. Tool that can assess the quality of a person's singing on the basis of acoustic parameters may prove useful. In order to determine parameters that could help evaluate the correctness of singing, a study was conducted on a group of 10 choir members and one professional singer. The study consisted of recording the singers' voice during singing and speech. The subjects performed simple vocal exercises consisting mainly in upward and downward sound modulation. In this study, portions of the recordings were analysed to determine parameters like Maximal Phonation Time (MPT), Singing Power Ratio (SPR) or signal integral. The values of obtained parameters for the choristers were compared with the results of the professional singer, which allowed to select those parameters that may be helpful in the evaluation of the singing voice. The parameters for which the connection between their value and singing correctness has been shown create a vector of features that can be used to assess the correctness of classical singing. The paper also describes further research plan.

Keywords: singing voice, voice analysis, acoustic processing, formants, singing pattern.

1. Introduction

Voice plays a fundamental role in human relationships and communication. Voice allows people to convey the information through words, emotions and expressions. In addition to fulfilling the communicative function in everyday life, human voice is also an instrument and work tool for singers, teachers, actors and many other professionals who intensively use voice in their work.

Phonation depends primarily on the correct functioning of the respiratory system and the phonatory organ, i.e. the larynx. Properly controlled muscles allow to produce a good quality sound without discomfort or excessive effort. Effective voice emission is an activity that involves the synchronization of physical processes such as breathing, phonation and resonance [1–3].

Singing is described as an extension of speech, however, the correct performance of this activity is a complex activity. Correct phonation requires hard work and training. In order to produce sound, whether during speech or singing, it is necessary to synchronize the correct functioning of the larynx and the vocal cords, breath correctly and engage resonant spaces.

The sound is formed in the larynx by the vocal cords. During the inhale diaphragm shrinks and increases the pressure of the expiratory air located below the glottis. Next a column of air is produced that extends the adducted and stretched vocal cords. The correct position of the vocal cords during phonation is shown in Fig. 1. At this stage, however, the sound is completely different from the one we hear. It does not have a color, volume or timbe, which are given to it by the resonators. The acoustic wave produced by the vocal cords is called laryngeal tone.

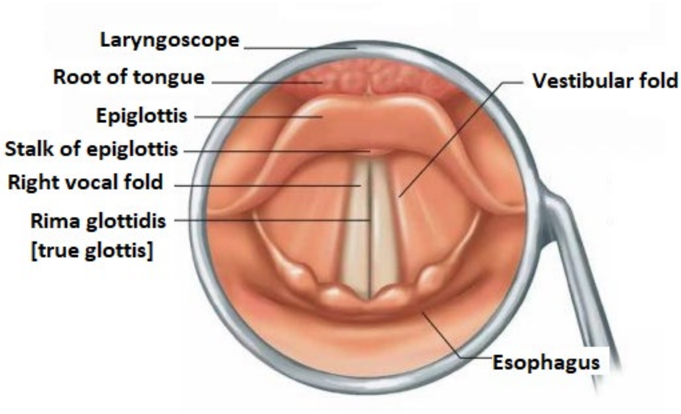

Figure 1. Vocal folds during correct phonation [4].

Vibrating vocal folds transmit vibrations to the laryngeal cartilages connected to them, then to the cartilages of the trachea and bronchi. Resonance is also created in resonators, i.e. air spaces that are limited by walls. In the human body, we can distinguish lower resonators (subglottis system, trachea, bronchi, chest) and upper resonators and these are: larynx, throat, mouth, nasal cavity and paranasal sinuses (Fig. 2). Among these anatomical structures, there are those that can change their arrangement and position, which has a huge impact on the process of correct voice emission, and these are: larynx, throat and oral cavity. The larynx is the first resonator, which in combination with subsequent air spaces (throat and mouth) forms a structure resembling a pipe. The larynx is able to change the capacity of this tube (vocal pathway) through its downward movement (with inspiration and downward movement of the mandible), backward (coupled with the shift of the tongue towards the front of the craniofacial region) or forward (accompanied by the elevation of the palate). This causes a change in the capacity of the resonators and, consequently, a change in the timbre of the sound. The oral cavity can change its shapes most effectively and by that allows to shape the timbre of the produced sound and strengthen its harmonic tones. In the classical singing technique, it is important to remember to use the upper resonators as they make the sound more powerful and give it more color [5].

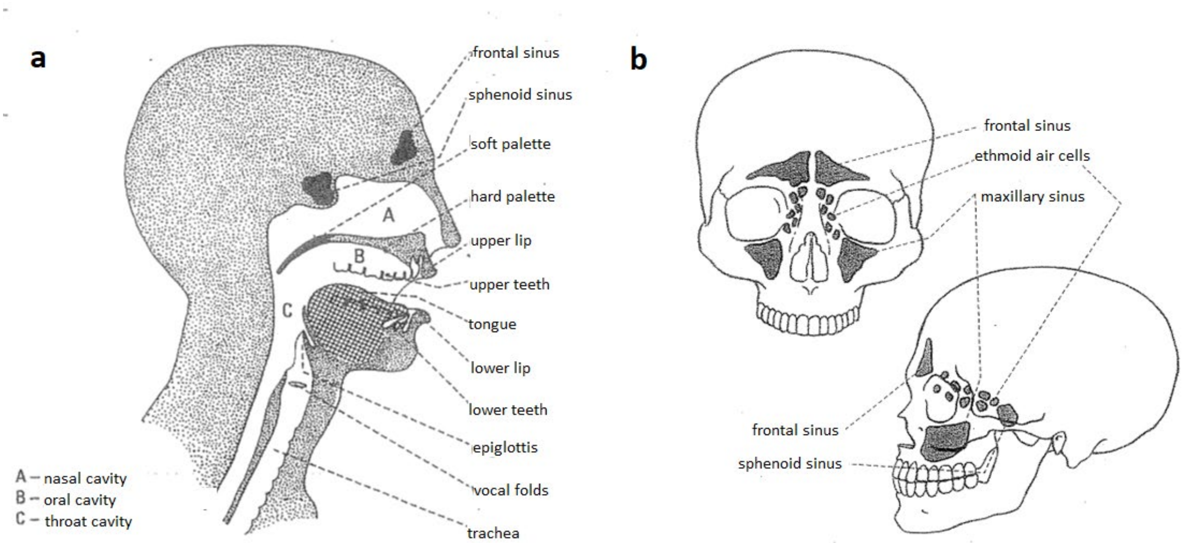

Figure 2. Upper resonator chambers (a) and paranasal sinuses detailed frontal and lateral view (b) [5].

The voices of singers are constantly influenced by external conditions, which is why these professionals are at risk of developing voice disorders and speech diseases. Thus, to avoid such results, singers should have knowledge of how and where the voice is created. An essential element is also to undergo proper vocal training and use proven techniques that will help in warming up the voice.

Unfortunately, some publications show that members of amateur choirs have problems with their voice and feel its fatigue even after performing vocal exercises before intense singing [6, 7]. Such observations may lead to the conclusion that choir members are not taught to properly use their voice and are not aware of how the sound is created in their voice apparatus. In this case the incorrect singing can be caused by wrong configuration of speech organs and this means wrong tongue position, inaccurate opening of the mouth or excessive tension of the laryngeal muscles. Any of these mistakes, if not eliminated, can cause an injury, which in turn can lead to many problems, diseases and serious pathologies of the voice organ such as acute laryngitis, laryngeal edema, polyps or even cancer. The negative effects of excessive voice strain have been assessed for different professions, each with its own characteristics [8–10].

The vocal folds are a very delicate part of the voice organ and therefore, it is important to properly educate choristers so that they could take control over their voice and operate it correctly. However, it is difficult to provide each choir member with the right care of vocal coach, because each person needs different amount of time to learn the correct setting of his or her phonatory organ. The solution to this problem may be to create an application or other tool that will be able to assess a person's singing in real time and indicate whether a given vocal exercise has been performed correctly or not. Such an assessment could be carried out by the application on the basis of the analysis of the recorded exercise and the calculation of the set acoustic parameters, the value of which could indicate the correctness of the performed exercise. The aim of this study was to find acoustic parameters that would be useful for the aforementioned assessment of the singing voice quality.

2. Methods

2.1. Acoustic signal recording

The singers' voice recordings were made using a high-quality SHURE BG 1.1 dynamic microphone combined with a ZOOM H1 digital voice recorder. All recordings have been saved in WAV format to obtain the highest possible sound quality. The recordings were sampled at a frequency of 48 kHz with a resolution of 16 bits.

2.2. Study group

The study was conducted on 11 people, including current and former members of the Con Fuoco Choir of the AGH University of Science and Technology in Krakow and one person who is a soloist of the Krakow Opera. Table 1 contains characteristics of individual people taking part in the study.

Table 1. Characterization of study participants: voice type, formation in singing, is the singer active, years of experience, sex (F – female, M – male).

Participant’s ID	Voice type	Formation in singing	Active singer	Years of singing	Sex
S1	soprano	Choir	Yes	10	F
S2	soprano	Choir	Yes	10	F
S3	soprano	Choir	Yes	5	F
S4	soprano	Opera singer	Yes	28	F
A1	alto	Choir	Yes	4	F
A2	alto	Choir	Yes	3	F
A3	alto	Choir	Yes	7	F
A4	alto	Former choir member	No (last active 12 months ago)	4	F
B1	baritone	Choir	Yes	3	M
B2	baritone	Choir	Yes	2	M
B3	baritone	Former choir member	No (last active 18 months ago)	7	M

Each of the former and current members of the choir remained under the care of a qualified voice emission teacher for several years, thanks to which they could increase control over their voice and the way it is produced. Due to her many years of experience and profession, a person with an S4 identifier was treated in the later part of the study as a model of correct voice emission.

2.3. Test exercises

Each of the participants took part in at least two solo recording sessions, during which different voice exercises were performed. Some of the exercises were performed by the study participants in two ways, that is, through the use of classical singing techniques and without them. All exercises performed by the singers are described in Tab. 2.

Table 2. List of exercises performed during a single recording session.

No.	Name of the exercise	Description
1	Maximum phonation time	Vowel ‘a’ sung at one pitch in a single breath.
2	Repetition of a simple melody	Short vocal exercise performed on vowels ‘a’,’e’,’i’,’o’ and ‘u’, climbing up the scale by halftones till the end of the singer’s range and then moving down the scale. The exercise was performed in two variants (with and without the use of classical singing techniques).
3	Vowels	Vowels sung for 2 seconds. The exercise was performed in two variants (with and without the use of classical singing techniques).
4	Intonation change	Change in intonation while maintaining a constant pitch. Singer starts from singing the vowel 'a' with using classical singing techniques and after about 2 seconds continues singing without using them.
5	Scale	Singing a scale on vowel ‘a’ going up by halftone until the participant’s end of range.
6	Glissando	Glide between lowest and highest pitch of participant’s voice range.

2.4. Description and parameterization of the acoustic signal

Obtained recordings were processed, including removing unnecessary fragments of recordings and selecting samples that will be useful for further analysis and calculations. Data pre-processing was performed using Audacity 2.4.2. The MATLAB environment (program version: R2020b) was used to process and analyze the received acoustic signals.

For the basic functional assessment of the respiratory system, the Maximum Phonation Time (MPT) was measured. It is a simple, but also highly valuable aerodynamic parameter, thanks to which the efficiency of the voice organ can be predetermined. The MPT value is around 20 seconds. Professional singers can achieve longer phonation times. It happens, however, that the obtained results will be shorter, which may be caused by care for the correctness of the emission [11]. The MPT measurement was performed in a standing, upright position. It was repeated three times and from the results obtained then the average value was calculated.

Recorded acoustic signals are discrete signals with a sampling frequency f_p, therefore consecutive time moments $x(n)$ were used to calculate the parameters in time domain such as: signal integral, mean value, energy, power and RMS (Root Mean Square). There was also Crest Factor calculated which is the ratio of max value of the signal and RMS. In addition, the Jitter and Shimmer characteristics of signal were also determined.

Samples of recordings were subjected to analysis in frequency domain using the STFT (Short-Time Fourier Transform), resulting in the time-frequency spectrum of the signal. STFT allowed then to prepare the spectrogram, which is defined as:

$$G_x(t,f) = |\mathrm{STFT}_x(t,f)|^2, \tag{1}$$

where t is the instantaneous value of time and f is value of frequency.
In practice, it is designated as:

$$G_x(t,f) = 20 \cdot \log_{10}(|\mathrm{STFT}_x(t,f)|^2). \tag{2}$$

During the analysis a Hamming window with a width of N = 1200 was used, which gave a time window resolution of 25 ms.

Assessing the sound quality can be also possible by formant analysis [12]. Formant is an aggregation of acoustic energy around a frequency in the speech wave. For correct way of singing there is so-called singing formant visible, which is third, fourth and fifth formant above the fundamental frequency grouping together (Fig. 3). Formant frequency can be described as:

$$F_{ij} = f_{ij} \Leftrightarrow \left[\left(\frac{\partial G(t_j,f)}{\partial f} = 0\right) \wedge \left(\frac{\partial^2 G(t_j,f)}{\partial f^2} < 0\right) \wedge \left(f_d \le F_{ij} \le f_g\right)\right], \tag{3}$$

where $G(t, f)$ is a spectrogram value, i is formant number, j is the moment in time (t) and f_d, f_g are scale cut-off frequencies.

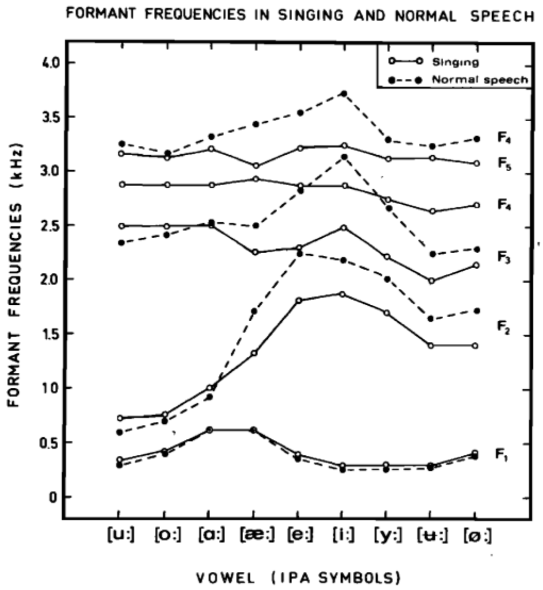

Figure 3. Average values of the formants for different vowels during normal speech (dashed lines) and during singing (solid lines). The F_4 during speaking has higher values than the F_5 formant during singing [12].

After determining the spectrum of the speech signal, $G_x(t, f)$ it was possible to determine the spectral moments, which are parameters used to describe the shape of the spectrum. The spectral moments can be defined from this general relation:

$$M_m(t) = \sum_{i=0}^{N} |G_x(t, f_i)| [f_i]^m, \tag{4}$$

where m is order, $G_x(t, f_i)$ is time-frequency spectrum at the time moment t, f_i – the middle frequency of the i band in frequency analysis, f_0, f_N are lower and upper frequency limiting the spectral moment determination band.

Using the calculated moments skewness (a measure of the asymmetry of the spectrum) and kurtosis (a measure of the flattening of the spectrum) were also defined. The next, and at the same time the last parameter determined based on the recordings is the Singing Power Ratio (SPR). This parameter is an indicator of the quality of the singing voice. SPR is calculated as the difference between the peak signal intensity (the highest value of the signal power spectrum) in the range of 2000 – 4000 Hz and its maximum value in the range 0 – 2000 Hz. It has been found that this measure can be used to differentiate the voices of experienced and untrained singers, as well as speech and singing recordings [13].

3. Results

Normalized and averaged values of the parameters for correct and incorrect phonation were prepared separately for group of women and men. The results of the person with S4 identifier were treated as a reference values for female singers. This paper will present the parameters which throughout the course of the study were found useful for voice performance evaluation.

Table 3 presents the average result of the measurement of the MPT for each singer, which was obtained as a result of exercise No. 1 (Tab. 2). For most of the subjects, the MPT oscillates around 20 seconds, which is the correct result [11]. The lowest score of 15 seconds was achieved by a person B3. This person already during the study reported difficulties with maintaining phonation for a longer time and feeling high tension of the abdominal muscles. It can be assumed that this effect was influenced by a long break in singing for the B3 person, which resulted in reduced control over the diaphragm and abdominal muscles. Abdominal muscle tension on inspiration causes excessive pressure on the diaphragm, causing discomfort and preventing proper phonation [14]. This discomfort contributes to faster contracting of the diaphragm, leading to the need for earlier exhalation and end of phonation.

Table 3. Average value of maximum phonation time for each singer.

Participant's ID	MPT[s]
S1	20
S2	18
S3	18
S4	19
A1	21
A2	20
A3	21
A4	17
B1	18
B2	18
B3	15

Best distinction between professional singer and choir members for parameters in time domain was found in the integral of the signal, its mean value, energy, power and RMS value. Values of all those parameters for choir members and professional singer were shown on the graphs in Fig. 4. Each graph compares the values of the parameters obtained for professional singer and for choir members during correct and wrong way of phonation. The values of other parameters in time domain – Crest Factor, Jitter and Shimmer, did not show a clear dependency on changing the way of singing.

The value of the integral of the signal (Fig. 4a) is higher when the sound is sung using the classical singing technique (choir_member_singing correctly vs choir_member_singing incorrectly). The same conclusion could be drawn from the comparison of values for the model and the femalechoir members for this parameter. Relationship between the sound correctness and parameters' values were found for mean value, power or RMS (respectively Figs. 4 b, d, e). Comparison of signal energy for choir members shown in Fig. 4c suggests a decrease in the value of this parameter as the quality of singing deteriorates. However for professional singer values of signal energy are even smaller then for the incorrect way of signing for other females. Presumably, this is due to the way experienced and amateur people sing. People with little experience, in this case choristers, tend to sing louder during vocal exercises, because they cannot properly adjust the power of their voice to the situation. The opposite is true for classical singers, who during exercises that do not require it, sing more quietly, focusing more on the way the sound is produced, which has a very good effect on the vocal apparatus, without straining it unnecessarily. For the recordings obtained during exercise No. 3, spectral moments, kurtosis and skewness were calculated, from which average values were determined for each of the recordings. After comparing values of the parameters in frequency domain for singing of the S4 person and the correct and incorrect singing of choristers, it was observed that only the spectral moment of the first order shows best distinction between professional singer and choir members (Fig. 4f) and therefore could be used to analyze the correctness and quality of singing. In addition, the SPR parameter was calculated, which for the professional has lower values then for untrained singer (Fig. 5) and can definitely be used to assess the quality of singing.

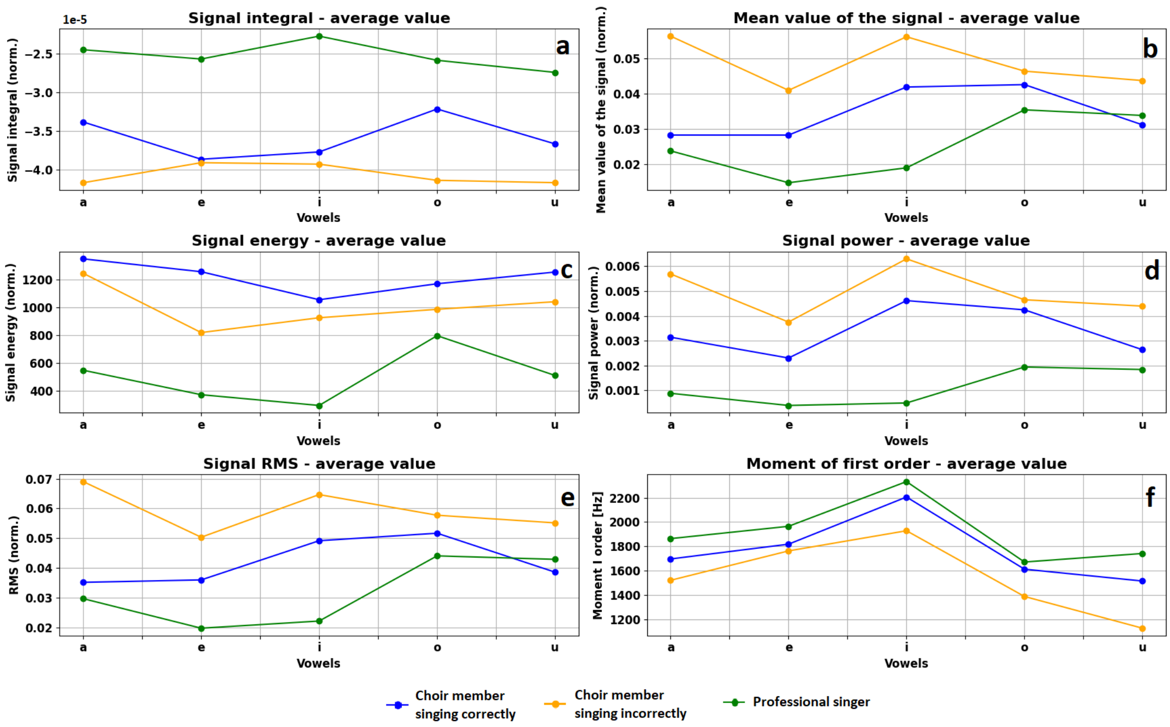

Figure 4. Comparison of parameters in time and frequency domain (signal integral, mean value, energy, power, RMS and spectral moment of I order) for female choir members singing correctly, incorrectly and for professional singer.

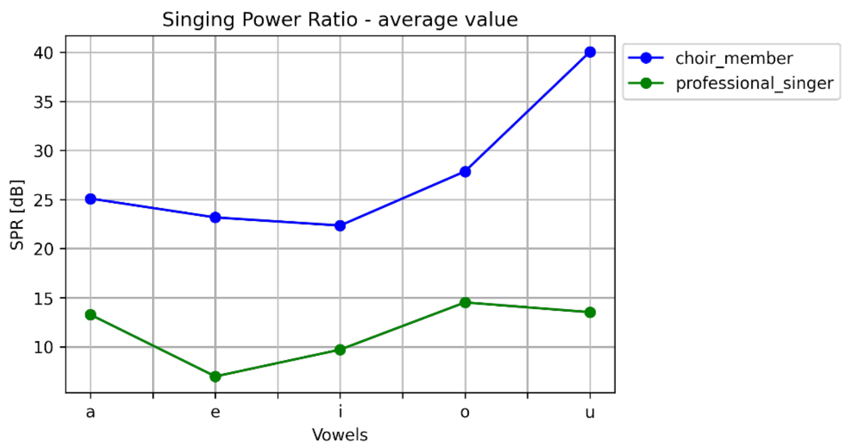

Figure 5. Comparison of the SPR parameter for the pattern and female choristers relative to the individual vowels sung.

For recordings obtained from exercise No. 2 (Tab. 2) spectrograms were prepared. The comparison of spectrograms of one choir member for correct and incorrect way of singing is presented in Fig. 6. can be noted that during correct singing (while maintaining classical singing techniques) higher formants are better visible on the spectrogram than during incorrect singing.

Using recordings from exercise No. 4, a spectrogram was prepared for each of them. An example result is shown in Fig. 7. As described in Tab. 2 the change in intonation occurred about 2 seconds after the start of phonation. For the B2 person, the transition from classical singing techniques to incorrect phonation occurs approximately 2.5 seconds after the start of singing. The change in intonation is very clearly visible on the spectrogram of the acoustic signal, where a decrease in the power of formants with frequencies above 2000 Hz can be observed. This can be also described as the disappearance of the singing formant.

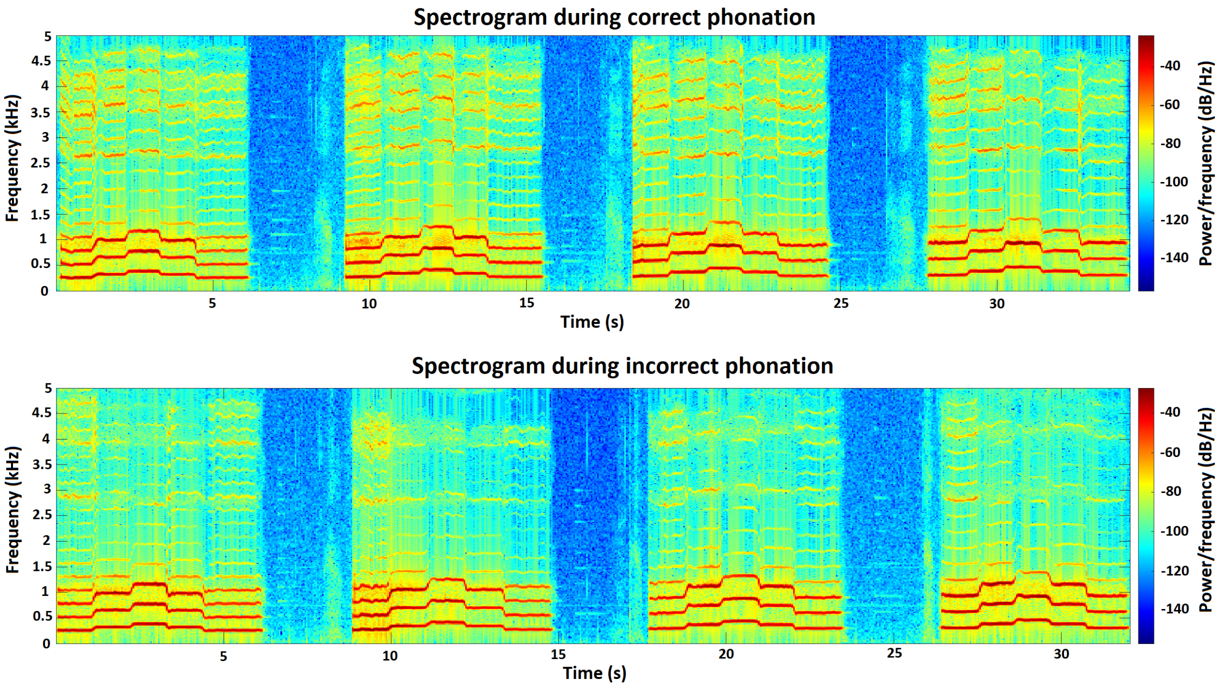

Figure 6. Spectrograms for a person with an A4 identifier during correct and incorrect singing the melody from exercise 2 (on the vowel 'o').

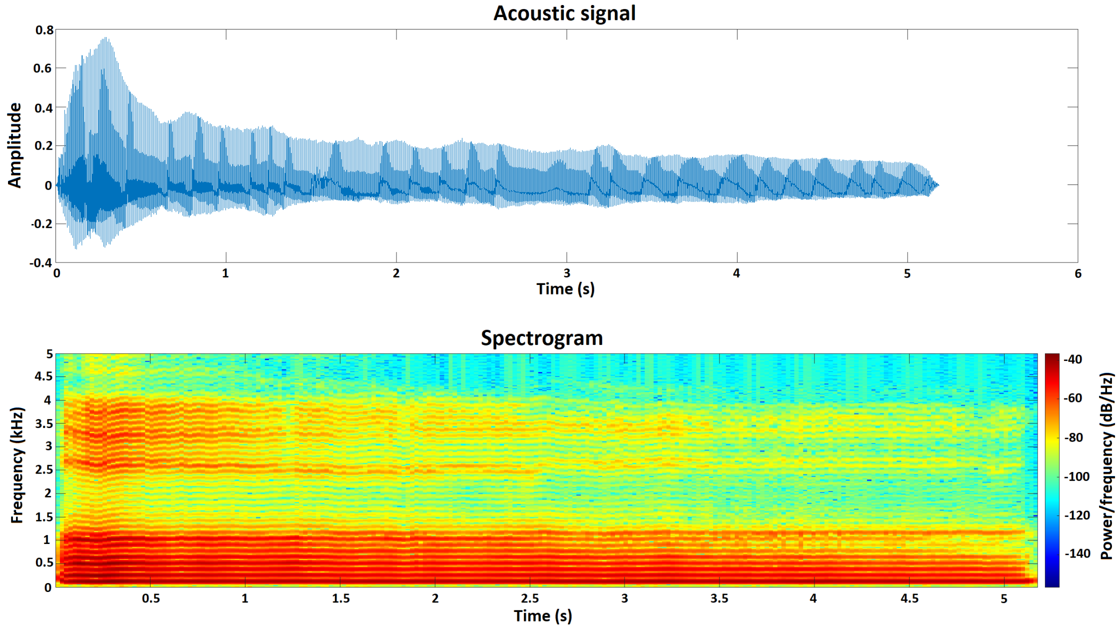

Figure 7. Signal waveform and spectrogram for a person with B2 id when changing intonation at a constant pitch.

As previously mentioned, during using the classical singing techniques there is singers formant visible on spectrograms. For each person the values of five formants were calculated and then used for evaluation of average values for groups of female and male amateur singers. In Fig. 8 formant values of each vowel of female choir group was set with results for professional singer. The singing formant is very clearly visible for the model of singing as grouping of formants F3, F4, and F5. The calculation of the formant frequencies definitely will be useful to determine the correctness and quality of classical singing.

The easiest way to parameterize this phenomenon is to calculate the distance between the F3 and F5 formant as the difference between these frequencies. Table 4 contains calculated F3 – F5 distances for the

S4 person – professional, female and male choir members. For the group of women and men, the values of this parameter were also given for incorrect singing, i.e. without the use of classical singing techniques. Analyzing the values in Tab. 4 we can easily see that the value of this parameter decreased when the participant used classical singing techniques. Similarly, a much smaller distance between F3 and F5 for the professional singer is observed.

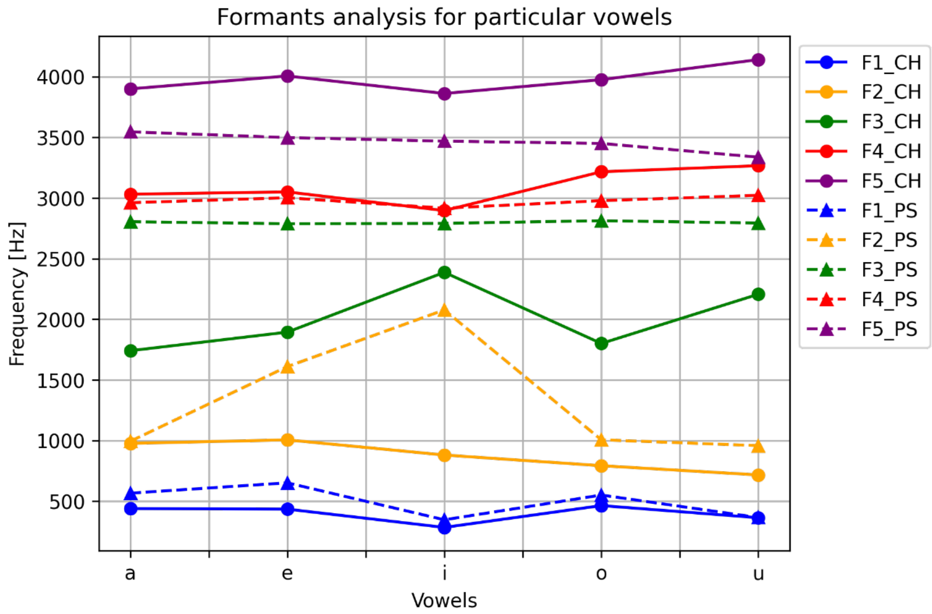

Figure 8. 1values of the formants for choir female members (Fx_CH) and the professional singer (Fx_PS). Formants of choir members are indicated by solid lines and circular markers, and formants of professional by dashed lines and triangular markers.

Table 4. The distance between formants F3 and F5 for the professional and the choristers when singing individual vowels.

Distance F3–F5 [Hz]					
Vowel	Professional	Women correctly	Women incorrectly	Men correctly	Men incorrectly
a	741.10	2159.05	2414.81	949.11	1455.08
e	710.10	2112.23	2241.10	1960.52	1444.36
i	678.00	1474.69	1491.39	1074.08	1329.86
o	637.10	2175.20	2813.89	1287.84	1874.25
u	543.20	1934.67	2496.16	981.05	1983.66

4. Conclusions

Voice is a crucial element of our lives. Human voice has foremost a communicative function but can be also considered as an instrument or work tool for people like actors, teachers or singers. It is said that singing is an extension of speech, however, it is a complex activity. Correct phonation requires hard work, many trainings, vocal and breathing exercises. Unfortunately publications [6, 7] and private observations show that members of amateur choirs often report voice fatigue or hoarseness, even if they performed warm-up vocal exercises before rehearsals or concerts. This may indicate a lack of accurate voice training and thus lead to frequent injuries and medical conditions of the phonatory organ. It is difficult to provide each member of the choir with sufficient care of a voice emission teacher. That is why it may be useful to develop a tool that could evaluate the quality of a person's singing on the basis of acoustic parameters and thus help in voice emission learning process.

Many organs are involved in the process of sound production. These include, among others: the larynx (the place where the sound is produced), the diaphragm and other organs of the respiratory system, as well as the frontal sinuses, wedge sinuses and other resonance chambers. Obtaining proper phonation during singing depends on ability to synchronize the proper functioning of the larynx (including the vocal cords), correct breathing and sufficient involvement of resonant spaces. Such a skill allows the singer to amplify the sound produced in the larynx and obtain appropriate timbre without excessive and unnecessary effort. Too much strain on vocal cords can be very harmful and can lead to the formation of diseases of the phonation system such as hoarseness, aphonia or singing nodules.

In order to determine the parameters that could be helpful in assessing the correctness of singing, research was carried out. It consisted of performing a few simple vocal exercises. The study was conducted on 11 people, including current and former choir members and one professional opera singer, which was treated as a reference person.

The recordings obtained as a result of the research were subjected to preliminary processing, selecting fragments needed for further analysis and calculations. The recordings of each participant were described in the time and frequency domains. For each of the obtained samples of recordings, many parameters were calculated, which are listed and described in Sect. 2. The results obtained for choristers were compared with the results of the reference person. Another comparison was done between the results of using classical singing techniques and without their use. Performed analysis shows that, unfortunately, not all of the examined parameters may be useful in assessing the singing of classical singers. Basic parameters such as voice scale or maximum phonation time (MPT) can be used to assess the progress of learning over time. Both voice and MPT scales should increase as you make progress in learning to sing. For the parameters analyzed in the time domain - integral, mean value, power and RMS of the signal were observed to show the best distinction between professional and unprofessional singer. For signal energy, Crest Factor, Jitter and Shimmer parameters no dependency was observed between their values and correct or incorrect way of phonation.

Analysis of the results obtained in the frequency domain confirmed that the so-called singing formant appears during singing (Fig. 7) and becomes more visible on the spectrogram when subjects used classical singing techniques (Fig. 6). In Fig. 8 clear grouping of formants F3, F4, and F5 was observed for the professional singer. The calculation of the formant frequencies definitely makes it possible to determine the correctness and quality of classical singing. The easiest way to parameterize this phenomenon is to calculate the distance between the F3 and F5 formants as the difference between these frequencies. Calculated distance decreased when the person taking part in the study was using classical singing techniques (Tab. 4). In frequency domain only parameter that may be useful in singing voice performance evaluation is the spectral moment of the first order. Comparison of the singing power factor (SPR) values for the choristers and the standard suggests that this parameter may also be useful in assessing the quality of singing.

The parameters for which the relation between their value and correct and incorrect singing has been shown create a vector of features that can be used to assess the correctness of classical singing. Many dependencies that can be noticed on the spectrograms have not yet been parameterized, so it may be helpful to use machine learning algorithms, which would be possible to create a pattern of correct singing, which could then be used to compare and evaluate singing at a given moment.

Learning to emit the voice is based on analyzing the quality of the sound sung by the student, assessing the activity of his abdominal muscles and tension of the muscles of the chest, neck and face. A voice emission teacher can only evaluate the correctness of phonation during the lesson, so a tool to support his work outside the classroom when the student is practicing alone would prove useful. In order for the student to receive feedback on the correctness of the exercises performed, it would be helpful to create a pattern of correct singing using machine learning. Such a formula could be obtained by recording several repetitions of a specific exercise under the supervision of the vocal coach who could confirm the correctness. On the basis of such classified recordings, parameters helpful in assessing the quality of singing would be calculated. In this way, a pattern would be created, which could then be used in real time to compare and assess the correctness of the exercise. Such a tool in the experienced hands of a professional can perfectly support the process of learning voice emission, which is extremely important for singers. Thanks to classical singing techniques, the singer gets less tired, the larynx does not make unnecessary movements, and this results in preventing vocal organ functions being impaired.

The research was carried out during the pandemic period, which caused difficulties in assembling a bigger, more diverse study group and taking more recordings, therefore it is proposed to carry out similar acoustic studies to those presented in this paper on a larger study group that will have representatives of each voice type. This will allow a more valid evaluation of the usefulness of the proposed parameters in determining the correctness of singing. In addition, it may be helpful to collect an electroglottographic signal that would indicate whether there will be tension in the neck and laryngeal muscles observed during singing for a particular individual. This could serve as a parameter to confirm the correctness of vocal emission, since muscle tension in this area should not occur during proper phonation. It is also worth taking video recordings during vocal exercises. The analysis of the video records could indicate the correctness of the posture adopted during voice emission and the position and width of the mouth opening.

This preliminary research has set up a background for the study that now is being conducted at the Department of Mechanics and Vibroacoustics at the AGH University of Science and Technology in Krakow, in cooperation with students of The Krzysztof Penderecki Music Academy in Krakow. The research will cover the study of the role and importance of acoustic processing of laryngeal (fundamental) tone in resonators in the singing voice. Present state of the art and knowledge tells modern singers or vocalists to sing with channeling the voice directly to the so-called "mask" [15–17], that is, to the front part of the facial cranium. Singing in mask means using the resonators withing the bones of the face to make the voice more powerful and obtain its appropriate timbre. However, the path of the acoustic wave that is produced by the vocal cords is longer and more complicated. A false perception of it or incorrect way of teaching may result in incorrect articulation of words, not achieving the full potential of the singer's voice or even in vocal organ function being impaired. Another subject of my thesis is also the reinterpretation of the role of the Eustachian tube in the processing of produced sound. To date, the tube is not included in scientific publications as an element that, consciously used, would allow better voice emission. It is only mentioned as a pressure equalizing element [18] and as a reason why we hear ourselves differently from our listeners when speaking [19].

This ongoing research will be a step forward to systematize and confirm the knowledge of the singing voice by studying the path of acoustic wave propagation through all elements of the vocal tract and the acoustic interpretation of resonators in the human voice. The results of this work can lead to changes in the science of voice emission and lead to the establishment of a method that is as effective as possible and at the same time the least effortless. Correct voice emission is important for singers, singing teachers and others who work with their voice, and can significantly reduce the likelihood of vocal tract fatigue and disease.

Acknowledgments

I am deeply grateful to my promoter, Professor Wiesław Wszołek for his guidance and support throughout my research.

Additional information

The author(s) declare: no competing financial interests and that all material taken from other sources (including their own published works) is clearly cited and that appropriate permits are obtained.

References

1. S.R. Kadiri, B. Yegnanarayana; Breathy to tense voice discrimination using zero-time windowing cepstral coefficients (ZTWCCs); Proceedings of the Annual Conference of the International Speech Communication Association, Hyderabad, India, September 2-8, 2018; INTERSPEECH, 2018, 232–236; DOI: 10.21437/Interspeech.2018-2498
2. S.R. Kadiri, P. Alku, B. Yegnanarayana; Analysis and classification of phonation types in speech and singing voice; Speech Commun, 2020, 118, 33–47; DOI: 10.1016/j.specom.2020.02.004
3. J.L. Gartner-Schmidt, D.F. Roth, T. G. Zullo, C. A. Rosen; Quantifying component parts of indirect and direct voice therapy related to different voice disorders; Journal of Voice, 2013, 27(2), 210–216; DOI: 10.1016/j.jvoice.2012.11.007
4. Krtań anatomia i patologie; https://altamed.pl/produkt/krtan-anatomia-i-patologie/ (in Polish, accessed on 2023.03.21)
5. B. Tarasiewicz; Mówię i śpiewam świadomie (in Polish), 1st ed.; TAiWPN UNIVERSITAS, 2003
6. F. Onofre, Y. de A. Prado, G.V.E. Rojas, D. M. Garcia, L. Aguiar-Ricz; Measurements of the Acoustic Speaking Voice After Vocal Warm-up and Cooldown in Choir Singers; Journal of Voice, 2017, 31(1), 129.e9-129.e14; DOI: 10.1016/j.jvoice.2015.12.004
7. D.L.P. Marchand, F.S. Kavaliunas, M. Cassol; The Effectiveness of the EASE Scale in the Development of a Vocal Warm-up Program for an Amateur Choir; Journal of Voice, 2019, 33(3), 310–316; DOI: 10.1016/j.jvoice.2017.11.003
8. M.Z. Lerner, B. Paskhover, L. Acton, N. Young; Voice disorders in actors; Journal of Voice, 2013, 27(6), 705–708; DOI: 10.1016/j.jvoice.2013.05.006
9. R.L. Middleton, V.A. Hinton; A Preliminary Investigation of the Vocal Behaviors and Characteristics of Female Pastors; Journal of Voice, 2009, 23(5), 594–602; DOI: 10.1016/j.jvoice.2008.01.002

10. A. Szabo Portela, B. Hammarberg, M. Södersten; Speaking fundamental frequency and phonation time during work and leisure time in vocally healthy preschool teachers measured with a voice accumulator; Folia Phoniatrica et Logopaedica, 2014, 65(2), 84–90; DOI: 10.1159/000354673
11. E. Sielska, K. Niemczyk; Diagnostic protocol in voice disorders; Polski Przegląd Otorynolaryngologiczny, 2015, 4(2), 12–19; DOI: 10.5604/20845308.1152198
12. J. Sundberg; Articulatory interpretation of the 'singing formant'; J. Acoust. Soc. Am., 1974, 55, 838–844; DOI: 10.1121/1.1914609
13. C. Watts, K. Barnes-Burroughs, J. Estis, D. Blanton; The singing power ratio as an objective measure of singing voice quality in untrained talented and nontalented singers; Journal of Voice, 2006, 20(1), 82–88; DOI: 10.1016/j.jvoice.2004.12.003
14. E. Kazanecka, A. Wrońska, A. Szkiełkowska, H. Skarżyński; The application of the Accent Method for the therapy of speech and singing disorders; Audiofonologia, 2004, 26, 81–84
15. C.H. Santoni; Influencing oral-nasal balance in speech and song; PhD Thesis, University of Toronto, Toronto, 2021
16. W.E. Ross; Importance of Good Technique in Singing; Music Educators Journal, 1961, 48(1), 91–95; DOI: 10.2307/3389724
17. I. Chatterjee, S. Mulia, S. K. Makar, V. Sinha; A Comparative Study among Trained Hindustani Classical Singers, Untrained Singers and Non-singers through Cepstral Analysis; International Journal of Phonosurgery & Laryngology, 2022, 12(1), 5–11; DOI: 10.5005/jp-journals-10023-1234
18. M. Evans; Vocal Qualities in Female Singing; PhD Thesis, University of York, York, 1995
19. D. Slater; Vocal physiology and the teaching of singing, 1st ed.; Foreman Press, 2009