DOI 10.24917/9788380840607.9

Ewa Brzdęk, Jolanta Zielińska, Wojciech Kulinowski

Laryngograph as a tool in diagnosis of voice and speech process disorders

Key words: electroglottography method, Laryngograph, speech signal diagnosis, voice disorders, voice profile, vocal cords irregularity

Abstract

The electroglottography method is an instrumental analysis using a device called Laryngograph, which is presented in a practical application in diagnosis issues of the speech signal for people with voice disorders. In this paper possibilities of this device are estimated. The visualization of the speech signal obtained with the use of Laryngograph allows detecting its acoustically and phonetically most important features and presenting them in a graphical form. The analysis process performed with the use of a computer and the presented computer attachment is easier, faster, and ensures higher quality than other methods. Computer voice recording enables not only visualization but also objective assessment and its repetitions. The aim of the research was the assessment and evaluation of the level of oral development of children with hearing impairments. The obtained results and following practical conclusions can serve as guidelines in clinical and speech therapy applications in voice rehabilitation and communication development.

Introduction

It is well known that the quality of communication depends on skills of understanding, speaking, and proper use of the communication system – the code. The success of oral communication mostly depends on the quality of voice and regularity of the phonation process. In present time, the way of oral communication, or, more precisely, the way of speaking of many people can hardly be called successful or correct. The reasons for this are disturbances of breathing, phonation, and articulation areas.

The aim of this paper is to extend and complement the information provided in two articles by J. Zielińska and E. Brzdęk (2015 a, b). However, this time we put stress on the difficulties in oral communication of some people with disability. We especially focus on two groups of disabled persons: people with paralysis and people with hearing impairments. The voice disorders of many of them make their communication significantly less efficient.

Voice disorders can be: organic, functional, or a combination of the two. Organic voice disorders fall into two groups: structural and neurogenic. Structural disorders occur in the situations when there is a dysfunction in the physical mechanism (this often concerns tissue or fluids of the vocal folds), which causes some lesion (physical abnormality) of the larynx. In this group we can distinguish the following disorders:

Contact Ulcers, Cysts, Granuloma, Haemorrhage, Hyperkeratosis, Laryngitis, Leukoplakia, Nodules (nodes), Papilloma, Polyps, Trauma, Miscellaneous growths (Guimares, Abberton 2005).

Neurogenic disorders are caused by problems in the nervous system as it interacts with the larynx. For example, we can observe the damage of the recurrent laryngeal nerve (which controls the movement of larynx) during cardiac, pulmonary, spinal and thyroid surgeries. When the nerve is damaged, it causes a paresis (weakness) or paralysis (complete lack of movement) in the vocal fold of the affected side. Other neurogenic voice disorders are related to other kinds of problems in the central nervous system. In this group we can distinguish the following disorders: Paralysis/Paresis, Spasmodic Dysphonia (Laryngeal Dystonia), Tremor (Benign Essential Tremor), and voice problems caused by another neurological disorder (e.g. Parkinson's Disease, Myasthenia Gravis, ALS/Lou Gherig's Disease).

A functional disorder means that the physical structure is normal, but the vocal mechanism is being used improperly or inefficiently, which is caused by poor muscle functioning. Thus, all functional disorders fall under the category of muscle tension dysphonia and correspond to various patterns of muscle tension. In this group we can distinguish the following disorders: Muscle Tension Dysphonia (general), Anterior-posterior Constriction, Hyper-abduction, Hyper-adduction, Pharyngeal Constriction, Ventricular Phonation, Vocal Fold Bowing.

The final category of functional voice disorder is the psychogenic disorder, in which poor voice quality becomes a symbolic, or outward, manifestation of some unresolved psychological conflict.

Psychogenic disorders exist because it is possible for the voice to be disturbed for psychological reasons. In this case, there is no structural reason for the voice disorder, and there may or may not be some pattern of muscle tension behind it. While it is quite common for a psychogenic component to exist in a voice disorder, voice disorders that are caused by a psychological disorder are relatively rare.

In general, there are difficulties with categorizing the voice disorders, because quite often different types of disorders interact (Zielińska, Brzdęk 2015 b, p. 34–35).

The problems connected with voice emission concern dynamic breathing, work of vocal cords, phonation, resonance, nasality, articulation, and prosodic features (stress, rhythm) of voice. They can be a cause or consequence of wrong emission and can be interrelated with different symptoms of organic and functional voice disorders. The important reasons for this are cerebral palsy (Lat. *paralysis cerebralis infantium*) and deafness (even partial) (Abberton 2000; Fourcin, Abberton 2008; Zielińska 2004; Grycman, Kaczmarek 2014).

The electroglottography method is one of the group methods of diagnosis of voice disorders. It is based on the acoustic analysis of voice samples, which allows the objective estimation of the phonatory function of the larynx.

The electroglottography method – general characteristics

Electroglottography is based on the electrical impedance measurement between vocal folds during their vibrations during the process of speaking. It is a function of their mutual position and it changes during larynx vibration. When vocal folds are cramped, the value of this function is smaller than when they are opened.

The signal obtained with the use of electroglottography, which represents the functioning of vocal folds, establishes the basis for Fx basic frequency determination (presented in Figure 1). Fx frequency is very important in the phonetic research.

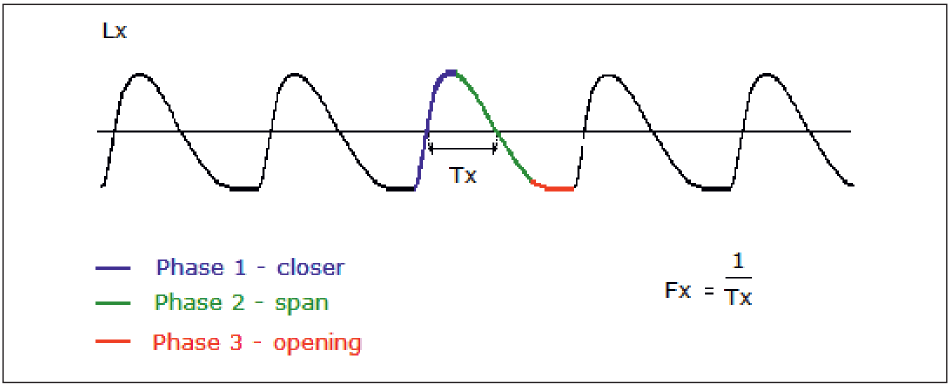

Figure 1. Sample curve of Lx signal, determination of basic Fx frequency [Zielińska 2004]

The initial part of the Lx signal of the functioning of vocal folds is the base for determination of the basic frequency Fx. Analysis of the signal cycle Lx helps to separate its three phases. The first phase is a quick signal growth, corresponding to the quick closing of the vocal folds. The second phase corresponds to the slightly slower signal fall and is connected with their opening. The third phase, which is a flat waveform, corresponds to the state when the vocal folds are open. Fx frequency is calculated while measuring each functioning cycle length in the middle of the Lx signal amplitude value (in practice, the measurement of time intervals between transitions of the signal through that level is conducted). The sought Fx value is the reverse of the Lx signal cycle length (Zielińska, Brzdęk 2015 a).

The electrical power diffused on the patient's throat is about 20 mW (when frequency is 1 MHz) and the top limit of the frequency is about 5 kHz. In the case of a normal male voice the ratio of the final output signal to the noise is about 40 dB. For small children relative noise is much bigger, however, successful results can be obtained also for new-born children.

The monitoring of vocal folds movements during normal speaking process takes advantage of this phenomenon. However, there are some disadvantages of the devices, which use it directly. Individual regulation for each patient is necessary and some distortions can appear, because the electrodes, which are used for impedance measurement, can translocate on the patient's throat. The technique, which presents an improved form of electroglottography and eliminates the mentioned disadvantages, is called "laryngographical" and was practically introduced in the Laryngograph device. The elimination of measurement errors was accomplished by:

introduction of electrodes with additional surrounding conductive rings, screened connections application, and usage of tri-pole net, of which the output is dependent on electrical conditions on the patient's throat only in the electrodes' direct neighbourhood

designing the circuit, which is dedicated to process the electrodes' output, and automatically compensates various impedances on the speaker's throat in such a way that the output is dependent only on fast vocal folds vibrations.

Possibilities of diagnosis of voice disorders by Laryngograph

Functioning of the computer attachment Laryngograph Processor PCLX is based on the electroglottography method. It is connected to a PC computer by means of an interface card, installed directly on IBM PC computer bus, which constitutes a coupler with the Laryngograph and software. This card contains a special system which processes signals and makes it possible to visualize them on the computer screen and record them. PCLX software called Speech Studio allows signals activation, visualization, on-line analysis, and storing them on the hard drive for later analysis and recovery.

Two electrodes are placed on both sides of the throat at the larynx level. The electrical impedance between them is the function of their mutual location, which is changed during larynx vibrations. It is smaller when vocal folds are close than when they are separated. The voltage on the speaker's neck is about 20MW (at the frequency of 1 MHz). The upper frequency limit is ca 5 kHz. In the case of a typical male voice the signal-to-noise ratio is ca 40 dB. In the case of small children and babies the relative noise is bigger, although satisfactory results can be obtained even for new-born babies. The computer screen shows vibrations of vocal folds and the basic frequency of this process.

The measuring result depends not only on vocal folds movements but also on larynx size and vibrating muscles. Thus, the whole phonic structure is evaluated. A separate part of the Laryngograph, namely a microphone, gives the signal which shows changes in the acoustic wave in time (line Sp) on the computer screen (presented in Figure 2).

These are oscillograms from which, through the analysis of visible changes of vocal wave amplitude in time, its basic acoustic properties can be derived, including sound and nasality. However, this way of presentation does

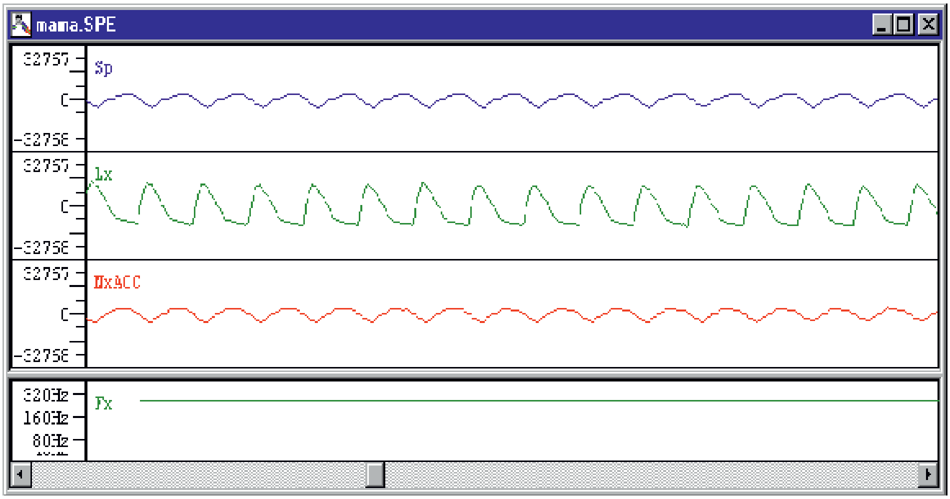

Figure 2. Computer presentation of signal 'mama', enabling the evaluation of the child's oral efficiency [Zielińska 2004]

not provide full information about the proper articulation of nasal phones by the examined person, the possible nasality process and its character, as well as the dynamics of air flow through the nose channel, which influences the quality, including the colour of the voice. All these possibilities are offered by the computer attachment called Nasality, which is a part of the research post-equipment. An electrode placed on nostrils is its inlet. The waveform generated by this electrode on the computer screen shows the dynamics of the air flow through the nose of the examined person, by measuring nostrils vibrations, like an accelerator which measures changes of speed in time, i.e. the acceleration of nostrils movement (line NxACC). Precise maximum amplitude values can be specified by creating a two-dimensional amplitude--frequency spectrum. Maximum values, visible on the computer screen, are called formants. Their frequency and amplitude characterize a phone and are closely related to articulator configuration that has appeared during its creation.

The initial signal of the laryngograph Lx (the second waveform in Figure 2) reflects the work of vocal folds and is the base for determination of the basic frequency Fx. Analysis of the signal cycle Lx helps to separate its three phases (waveform in Figure 1).

The sought Fx value is the reverse of Lx signal cycle length. The shape of Lx signal gives information about any abnormality of vocal tract work, especially the work of vocal folds. Sp signal (the first waveform in Figure 2), which is obtained from the microphone inlet, enables a dynamic visualization of signal timing, evaluation of its volume, loudness and exercising dentalised phones, e.g. Polish pairs “s – sz” (/s/ - / /), “z- rz” (/z/ - / /), which are very often misused and mistaken by persons having oral problems, including lisping. The analysis of Fx signal provides information about all parts of the proper signal, which are easily evaluated even by a child. They are: proper dynamic breathing, economical breathing, work of vocal chords, proper articulation, prosodic features, such as: pace, rhythm, stress, or intonation, i.e. these elements which decide about signal expression and, additionally, show the level of phone voicing (no voicing – no picture). NxACC signal (third waveform in Figure 2) enables determining proper nasal phones production, including the work of the glottis. Such analysis of voice parameters enables drawing many practical diagnostic-therapeutic conclusions (Baken 1987; Herbst, Fitch, Svec 2010).

Quantitative analysis provides a range of connected speech and voice analyses, which is only available from Laryngograph. It is important clinically and in voice research. Some concepts that are introduced can be new to many users, so groups of analyses have been set up to guide through QA. The analysis groups are numbered in a way that provides a basis for structured training and goes from simple pitch measurements to complex multi-component ones (Abberton, Howard, Fourcin 1996):

1. Frequency Fx — pitch range and regularity.
2. Amplitude Ax — loudness range and regularity.
3. Contact Qx — percentage range and regularity.
4. Loudness and contact and frequency.
5. Pitch and Contact.
6. Frequency Fx and contact Qx — pitch and quality.
7. Aspects of Voice Quality.
8. Frequency Related Measurements.
9. Elapsed Time Measurements.

The prototype of the presented device was introduced for the first time at the Faculty of Phonetics of the University College in London (Abberton, Howard, Fourcin 1996). Nowadays it is produced in Laryngograph Ltd. Company. During several years it has been put into practice in such domains as:

- speech therapy, in order to improve speech processes,
- clinical examination of nose and throat, to diagnose the work of vocal folds after surgical procedures,
- improvement of singers' and announcers' voice,
- examination of relations between speech and work of the vocal cords,
- teaching foreign languages.

Practical usage of the laryngograph in evaluation of speech

The most important element of the speech signal analysis is the way the signal is sampled and remembered by the computer. Computer attachment Laryngograph Processor PCLX possesses three input signal detectors. These are copper electrodes placed on both sides of the patient's throat using an adjustable band. A necessary condition of acquiring correct research results is the way that electrodes are placed on the patient's throat. It depends on the age of the person: if the person is younger, electrodes should be placed higher, assuming the central location between the 4th and 7th neck vertebra in the case of adults. This fact results directly from the physiological structure of the larynx. Fast growth of the larynx takes place in the case of every child until the age of 5. The second period of growth falls on puberty age. The larynx structure is changing: it is growing horizontally, especially in the case of boys, and it is associated with the voice mutation. This process takes place when children are 13 -15 years old, but sometimes lasts longer, till 19 years of age. In this period the neck gets longer and the larynx moves down. At that age it is important to pay special attention to physiologically correct placement of the laryngograph electrodes. The second decoder (microphone) should be placed about 2.5cm from the patient's mouth (Zielińska 2004).

During the voice generation process three basic phases take place and they can be fully monitored and analysed by computer software called Speech Studio. These three phases are: generation of stimulated vibrations, their modification in the vocal resonator, and radiation. The signal which is generated in the larynx during voiced sound pronunciation, by vocal cords in the role of stimulation generator, is a periodic flow, characterized by the basic frequency Fx, which was described in details earlier (Figure 1). In the narrowness of the voice channel the signal is modulated. The adjustment of the voice tract, i.e. articulators configuration, is characterized by formants, which are defined based on the average amplitude-frequency spectrum (presented in Figure 4).

Figure 3 presents the main screen of the integrated system for voice examination Speech Studio, which has all the possibilities of speech signal analysis that have been described previously. During work with this program there are two toolbars available (top and side one), which allows to perform several op-

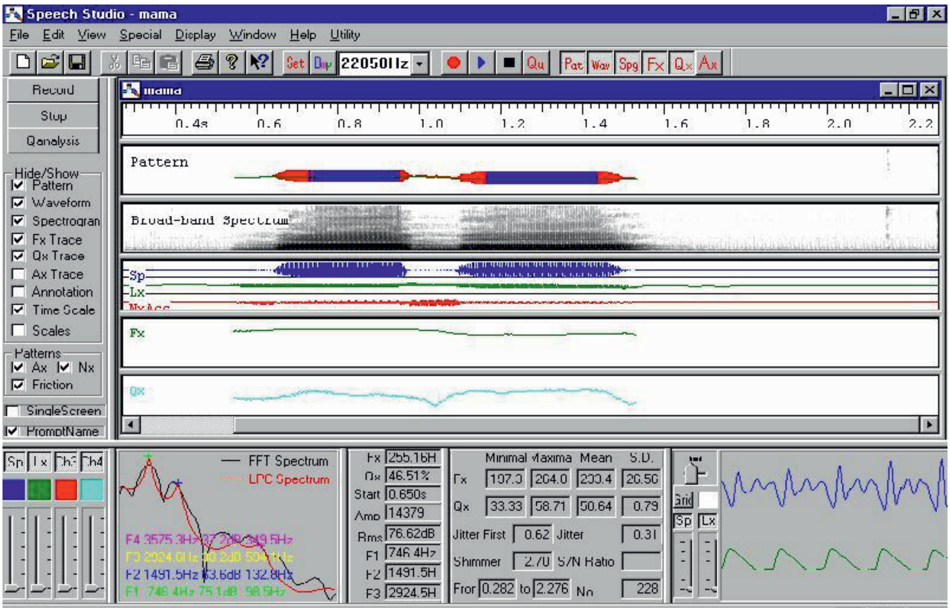

Figure 3. Integrated system for voice examination Speech Studio [Zielińska 2004]

erations, starting with signal activation, its remembering, recovering, analysis, to configuring the amount and type of visible flows and scales.

The fundamental part of the screen is seven flows, which characterize the speech signal. The two most important flows present changes of amplitude in time and the work of vocal folds. These charts can be enlarged in the bottom right window and allow an immediate preliminary diagnosis of the voice pathology and correct adjustment of equipment parameters. Signals amplification can be regulated by attachment knobs or sliders placed in the bottom left corner of the screen. In the bottom part of the screen there are also two other windows. In the first one average amplitude-frequency spectrum is displayed. In the second one there are values which characterize the speech profile of the patient. In this profile the following parameters are defined: minimum, maximum, and average value, standard deviation for the basic Fx frequency and occlusion factor Qx (WZG), and also other parameters describing disturbances of the frequency (Jitter) and amplitude (Shimmer). The flow of the occlusion factor Qx depends directly on the Lx flow and is presented as a percentage value. It determines the percentage usage of time when vocal folds are opened, that is the time of occlusion and opening during utterance pronunciation (the sum of phase 1 and phase 2 durations in Figure 1).

During this type of research, related to the speech signal analysis, on the computer screen many flows are simultaneously presented, characterizing the speech signal. On the basis of these flows we can define both the numerical parameters describing the voice during appropriate analyses, and qualitative conclusions. This program allows reducing the number of presented charts to the selected ones.

The obtained result depends not only on vocal folds movements, but also on the size of the larynx and on the weight of vibrating muscles. The signal, which is labelled on the computer screen as Lx, represents changes of the acoustic wave pressure. This is an oscillogram, and based on it, some initial information about the structural and functional pathological changes of the larynx can be obtained (Rothenberg, Mahshie 1988).

When the signal is more jagged and irregular, then physiological changes within the diagnosed organ are bigger. During the analysis of the presented changes of vocal wave amplitude it is possible to determine its basic acoustical features, sonority and nasality. The equality of random and determinate elements and the connection between individual and semantic aspects is clearly revealed in it (Rothenberg, 1992).

It is the main reason why time flows cannot be used in automatic speech recognition (ARM), however, they are widely used in acoustical features of speech research, especially the outline of signal amplitude. It is a widely examined issue, which is described in many works concerning phonetics.

The second chart presented in Figure 3 is called spectrograph. It is a broadband spectrum, which is used in acoustical evaluation of normal and pathological phonation. So it determines one of the phonetic parameters of speech at the segment level. It also allows differentiating the opposition: sonority – soundlessness, and evaluating prosodic elements of pronunciation, especially intonation. The flow presents changes of amplitude in time and its level, so the signal amplitude specified by the degree of darkness.

The next spectrum which is displayed on the computer screen (Figure 3) is called Sp and allows to visualize the time-based pronunciation flows, to define the degree of voice intensity, and to estimate the volume of sound. However, properly performed analysis of the basic Fx frequency gives information about all elements of correct pronunciation. These are: dynamic breathing, economical breathing, work of the vocal folds, pitch of the voice, prosodic elements of the utterance, such as: speed, rhythm, stress, and additional presentation of sound resonance degree (if there is no resonance, there is also no image on the screen). In the case of analysing correct articulation, mainly the shape of the larynx – mouth vocal channel, where the shape is modulated, the amplitude-frequency spectrum of the utterance is very useful (Figure 4).

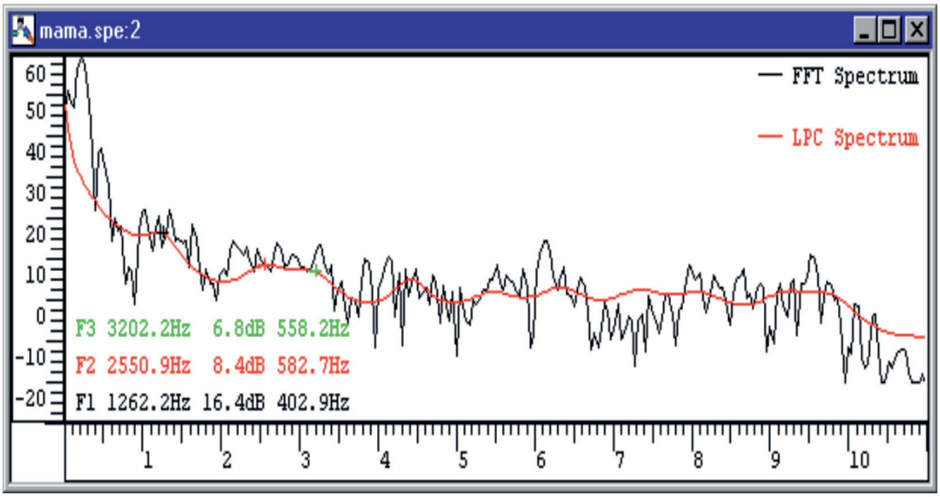

Figure 4. The average amplitude-frequency spectrum, “mama” utterance, Speech Studio program [Zielińska 2004]

The presented speech signal spectrum is obtained with appropriate computer procedures, which are based on the FFT algorithm (Fast Fourier Transformation) and LPC (linear prediction algorithm). It allows to determine formants in selected periods of time and the momentary spectrum in specific points of time. For each format its maximum frequency value, amplitude, and frequency band is presented (Herbst, Fitch, Svec 2010).

There are various possibilities of research material analysis, starting with statistical calculations, parametrical calculations, to graphical processing. In order to obtain a more precise signal view, we can select any part of it and freely enlarge, so that even the smallest pathological change can be detected. This software also enables measurements of time intervals between individual points of the signal. It can be performed by appropriate placing of two cursors on the screen, and time interval between them will be measured.

The Fx signal analysis is based on determinations of 23 histograms, including the CFx histogram, which presents the vocal cords work irregularity (Figure 5).

Computer analysis of speech signal allows obtaining the voice profile of the patient in the form of acoustic parameters (Figure 6).

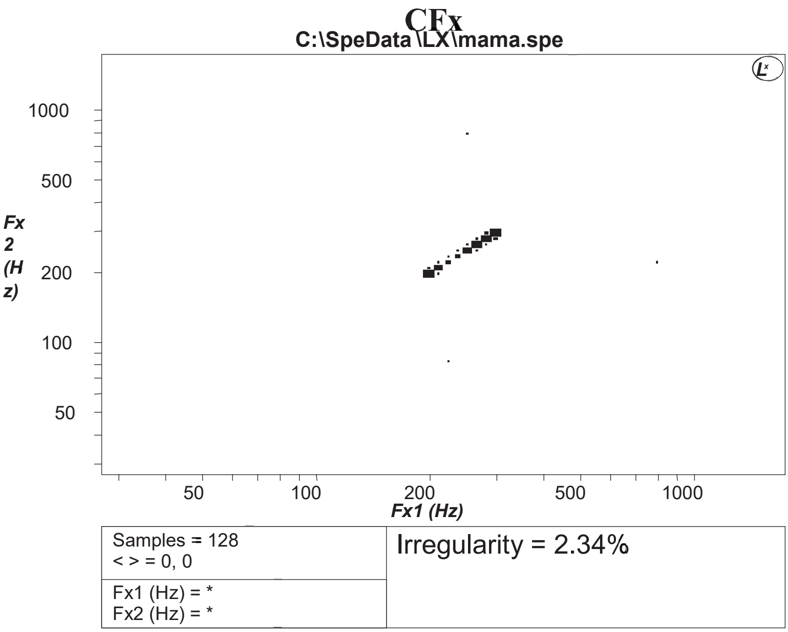

Figure 5. Sample CFx histogram, which presents the vocal cords work irregularity [Zielińska 2004]

nama.spe:2

Report – Vocal Profile Analysis

No. Samples	8	Duration	0.714 s
Minimal Fx	194.43 Hz	Maximal Fx	211.95 Hz
Average Fx	203.16 Hz	S.D. Fx	5.08 Hz
Minimal Qx	36.28 %	Maximal Qx	47.27 %
Average Qx	42.47 %	S.D. Qx	2.19 %
Jitter First	2.16 %	Jitter Second	0.64 %
Shimmer+	24.34 %	Shimmer-	-15.69 %

Figure 6. Sample voice profile for "mama" utterance, Speech Studio program [Zielińska 2004]

The voice profile contains the average value of basic Fx frequency, its minimum and maximum value. It also enables determining frequency oscillations, standard deviation, the factor of frequency disturbance called Jitter, percentage value of the average occlusion factor Qx, its minimum and maximum value and standard deviation. Additionally, as the last information, the value of voice amplitude disturbance factor (called Shimmer) is presented (Abberton, Carlson 1999; Zielińska 2004; Zielińska, Brzdęk 2015 a).

The experimental research was conducted in Cracow (Poland) by J. Zielińska (2004). This investigation concerned the assessment and evaluation of changes in the level of oral development of children with hearing impairments. The obtained results and following practical conclusions can serve as guidelines in clinical and speech therapy applications in voice rehabilitation and communication development of groups of deaf persons and cerebral palsy patients.

Computer analysis of selected children's speech samples with significant or deep impairment of hearing proves that the majority of them show symptoms characteristic of hyperfunctional dysphonia with mixed or closed nasality. Nasality is functional and is caused by the improper work of the soft palate and glottis. The tests showed: improper, mostly collar bone – rib breathing channel, harsh voice, too low or to high, generally with smaller sonority. Also there appeared: shorter phonation time, disturbed prosodic features, and voice tone with nasality features. The tests and the following statistic conclusions show a very high efficiency of the diagnostic-rehabilitation method suggested here.

The tests of the oral ability of children with significant or deep hearing impairment enabled specification of the pathology of changes characteristic of this group in some selected evaluation categories and gave preliminary possibilities of removing them on the basis of mainly voice breathing rehabilitation connected with the release of larynx muscle tonus and the work of soft palate muscles and glottis. The fact that the research was conducted on a statistically valid group of children gave the possibility to generalize the presented conclusions and to compare them to clinical tests performed in this subject area. The analysis of the efficiency of the diagnostic-therapeutic procedure showed that the therapy aiming at achieving proper dynamic breathing is connected with the work on nasality and prosodic features of speech, while the work of vocal cords is connected with the proper voice pitch. The research gave unique physiological results. Dynamic breathing, the work of vocal cords, voice pitch, and dynamic breathing with nasality change according to the rules of mathematical implication, which was confirmed by practical tests on a statistically valid group of children. For example, if dynamic breathing is proper, then the work of vocal cords must be proper, too, whereas improper breathing does not have to cause the improper work of vocal cords.

The experimental research tests filled the methodology and application gap in the process of using computer technology for diagnosis and development of oral efficiency in deaf children's hearing system. Their results and conclusions can serve as guidelines in clinical and speech therapy applications in voice rehabilitation and communication in Polish (Zielińska 2004).

Conclusions

The electroglottography method allows a thorough analysis of the speech signal, saved in the computer memory. For example, it can be a general spectral analysis or a detailed analysis based on formants. By means of basic Fx frequency visualization on the computer screen it is possible to estimate the voice pitch as well as to obtain the full acoustic profile of voice, with the use of instrumental analysis. The general conclusion is that the methods of speech signal analysis can be as complex as the examined signal itself.

The scope of usage of the electroglottography method, presented in this paper, depends mainly on the designated research goal and competence of the person performing the examination. The undisputed fact is that the presented possibilities are really extensive and, what is the most important fact, comparatively easy to obtain and interpret.

References

Abberton E., Howard D., Fourcin A., *Laryngograph assessment of normal voice: a tutorial,* Clinical Linguistic and Phonetics 3, Informa Healthcare 1996, p. 243–259.

Abberton E., *Voice quality of deaf speakers,* (in:) Kent, R. D., Ball, M. J. (eds.) Voice quality measurement, Singular Publishing 2000, p. 449–460.

Abberton E., Carlson E., *A window on the voice - how I use computers in therapy,* Speech and Language Therapy in Practice, 1999, p. 226–227.

Baken R. J., *Clinical measurement of speech and voice,* Boston: College Hill Press, 1987.

Fourcin A., Abberton E., *Hearing and phonetic criteria in voice measurement: clinical applications*, Logopedics Phoniatrics Vocology *33*(1/2008), p. 35–48.

Grycman M., Kaczmarek B. B., *Podręczny słownik terminów AAC (komunikacji wspomagającej i alternatywnej),* Impuls, Kraków 2014.

Guimaraes I., Abberton E., *Health and voice quality in smokers: an exploratory investigation,* Logopedics Phoniatrics Vocology 30(3), 2005, p. 85–191.

Herbst C. T., Fitch W. T., Svec J. G., *Electroglottographic wavegrams: a technique for visualizing vocal fold dynamics noninvasively,* J. Acoust. Soc. Am. 128, 2010, p. 3070–3078.

Rothenberg M., *A multichannel electroglottograph,* Journal of Voice 6, 1992, p. 36–43.

Rothenberg M., Mahshie J.J., Monitoring vocal fold abduction through vocal fold contact area, *J. Speech Hear Res.* 31, 1988, p. 338–351.

Zielińska J., *Diagnoza i terapia sprawności ortofonicznej dzieci z uszkodzeniem słuchu wspomagane techniką komputerową,* Wyd. Naukowe AP, Kraków 2004.

Zielińska J., Brzdęk E., The application of laryngograph in research of quality of speech signal: the electroglotography method, Recent Researches in Electrical and Computer Engineering, WSEAS Press 2015 a, p. 30–35.

Zielińska J., Brzdęk E., Laryngograph as a tool in diagnosis and therapy of voice disorders, Recent Advances of Biology and Biomedicine Series 9, WSEAS Press 2015 b, p. 34–38.