

SELECTED METHODS OF PATHOLOGICAL SPEECH SIGNAL ANALYSIS

W. WSZOLEK

AGH University of Science and Technology
Faculty of Mechanical Engineering and Robotics
Department of Mechanics and Vibroacoustics
Al. Mickiewicza 30, 30-059 Kraków, Poland
e-mail: wwszolek@agh.edu.pl

(received June 15, 2006; accepted September 30, 2006)

Selected results of examinations, having been performed by the author for several years, concerning the evaluation of selected methods of transforming signal in analysis tasks and pathological speech evaluation usability were presented in the article. In many issues of medical diagnosis, as well as in the planning of some illnesses therapy and rehabilitation, it is necessary to evaluate the signal of deformed speech. This evaluation can concern evaluation of the signal deformation degree and in this case the task is to present in a quantitative dimension (preferably scalar) the measure of deviation between the measured pathological speech signal and the abstract signal, which can be acknowledged as a pattern of the correct speech. Possessing such a scalar measurement, of the considered signal deformation degree, enables to monitor the illness process development or remission, which has a key importance in the current monitoring of the therapy effects or/and many illnesses rehabilitation. Pathological speech evaluation can also go into the direction of a classification and determination of a type of the analysed signal, which can, in some cases, have a direct connection with anatomical and functional causes as well as conditions of the considered illness. The classification of pathological speech signals can facilitate the diagnosis by pointing the most probable causes of the speech signal pathological deformation. Such a classification can also help in the optimal therapy selection and in determining rehabilitation recommendations.

The methods of transformation, analysis, classification and speech signal recognition have been known apparently for many years, as with no difficulty, many writings discussing these concepts and presenting results concerning both basic examinations results and many application works, can be found. The problem, presented in this article, is strongly different from the majority of works, which were published by other authors, as most of well known works, concerning the speech signal analysis (etc.), are directed to understand the content of the statement (automatic speech understanding, a speech-writing conversion, steering of the devices by speech signal), or relatively by determining the identity of the speaker (automatic identification or authentication of the speaker). Meanwhile, in the pathological speech analysis the semantic statement content is not essential; it is also not important who the speaker is. The research subject is the speech articulation process itself and all its pathological deformations, which determines both the used signal analysis tools as well as the techniques of the selected objects recognition, which are the forms of the particular ill person speech deformation forms in comparison to the speech of the whole sound people population [5].

Key words: speech analysis, speech processing, speech recognition, pathological speech, automatic diagnostics, computer based therapy monitoring, biomedical engineering.

1. Introduction

Speech acoustics provides several methods of speech signal quality evaluation, enabling a multilateral analysis with its results visualisation and their changeability process during speaking [7]. However, the direct analysis of this type of a process is very complex and requires a lot of experience, especially in case of pathological speech analysis. Hence the methods of automation analysis processes and speech signals recognition are developed and still enriched, and the results of these examinations are presented in number of works, including this article [2, 3]. Because the bibliography of computer analysis transformation methods field and recognition of various sound pictures is very rich, and the number of works, which appeared all over the world in the context of speech recognition problem can be estimated as tens of thousands positions, before we go to the detailed consideration it is necessary to point what new and original values this particular article implements, in relation to this extremely rich and diversified bibliography of the subject [4, 8, 11, 12]. Namely, the key term is the fact that the subject of research in this work is a pathological speech signal, and the crucial aim of the research is determining the nature, kind, and degree of illness changes advance, manifested by acoustic changes in the considered speech signal.

Years of experience of the author and people, with whom he has performed the research for several years prove, that surprisingly in many medical diagnosis issues, and in planning and monitoring of speech organs or the organs connected with speech (e.g. stomatologic prosthetics) therapy and rehabilitation, it is necessary to evaluate the deformed speech signal quality [9, 13, 14]. In the article selected results of years of the author's research concerning applying modernised methods of transforming acoustic signals in several selected medical problems were presented as syntactic conclusions. In the initial research, performed few years ago, it was proved that in the issues connected with analysis, evaluation and classification of the pathological speech signal, the standard methods of the signal transformation and classification, used in the semantic speech recognition (understanding the statement content) or in the voices recognition (speaker identification) fail completely [12, 14].

In a typical research, concerning speech recognition, the aim is mostly to disclose (through selected parameters) the semantic aspects of the pronounced text, in case of medical diagnosis, based on the pathologically deformed speech signal analysis; the semantic content of the statement is insignificant (and can even be treated as a disturbance). Based on the earlier research, performed by the author himself and in mixed research teams, including the acousticians and doctors of various specialisations, it was stated, that in these issues the standard speech signal transformation and classification methods, used for example in the semantic speech content or the voices recognition fail completely [10, 16]. As well the standard techniques of the speech signal parameterisation, such as linear prediction coefficients or cepstral coefficients cannot satisfactorily describe the specifics of pathological speech features, for the purpose of its different phonetic and acoustic structure in relation to the correct speech signal, and also because of the reason, that the aim of recognition in this case is completely different [4, 18]. Se-

lected methods of transformation and signal analysis in the task of pathological speech evaluation were presented in the article; moreover their usefulness in the selected medical issues was discussed. The wide range of various medical issues, which are presented in the work as the problem areas of the performed research are worth stressing. It proves that the acoustic techniques and methods described in the article, as well as the techniques related to transforming signals, are quite universal, although from the medical point of view quite often they can concern very different tasks.

It was stated that the most important (and the most difficult!) element of research work preceding a practical using speech as the source of medically useful diagnostic and prognostic information, is finding and describing these signal parameters, which are maximally independent, both from the context and the personal features of the examined voice. Additionally searched features of the signal must be maximally sensitive to its deformations in this layer, even the small ones, which is connected with the structure and functioning of speech signal generators (larynx and strangulations, being the source of speech noise constituents) and with the voice tract structure being used during articulation. During the research a particular attention was put to the analysis and description of the space structure features describing pathological speech signal, because accurate knowledge of the space features topology (uneasy for the direct evaluation because of the multi-space character) enables the later effective applying appropriate methods of automatic [2, 13].

In the present work the main attention was put to the new scientific results, acquired during the author's personal research, that is why the already well described, in specialist writings, methods of pathological speech analysis, for example included in the MDVP (Multi-Dimensional Voice Program) of the KAY company, were omitted. This system offers, in a convenient form, the possibilities of performing many analysis directed towards the pathological speech specifics, and that is why it is very attractive for doctors (phoniatries, laryngologists), regardless possessing few limitations (e.g. measurement accuracy F_0). However, appreciating, in the positive sense, the possibilities, which are given by the MDVP system, it can be stated that, on one hand these methods have already been well known, so researching them does not provide new values in the scientific aspect, and on the other hand the pathological speech signal parameterisation, offered by this system, is not sufficient for solving many practical tasks, which we encountered in our researches [13]. That is why this issue was chosen as a research topic and several new results were obtained, which will be discussed below briefly.

2. Research assumptions

The point of departure in the described research was a statement that the process of recognition and evaluation of acoustic deformation of pathological speech signal means assigning the obtained, as a result of research, the acoustic pictures to particular classes; however in distinction from the other diagnostic tasks we encounter a situation in which counting set of classes is usually unknown. It is worth to pay attention to

this detail and to stress its importance: in the statement recognition, in order to, for example, operate particular machines or devices by voice; set of recognizable classes of pictures is determined in advance. The automat is to recognize which of the previously planned comments were pronounced, or should certify, that the registered sound signal does not belong to any of the remembered patterns, which means that either it is not a speech signal at all or a casual statement was caught, the statement which is an unknown command.

Similarly, in the tasks of speaker's recognition we also deal with the already determined set of patterns. We have to either state, which authorised person pronounced the phrase, enabling to identify their voice (of course, paying particular attention to the possibility that the speaker comes from the set of people who are authorised to use a particular resource, and then a special alarm procedure should be activated), or we already determine a particular hypothesis (that a speaker is a particular person) and biometric methods have the task to verify this statement in the 'true or false' categories. However, in the pathological speech signal deformation recognition, there are no fixed distinguishers of deformed speech pictures classes, and moreover there is no possibility to foresee in advance how many classes it will be possible to distinguish. It can be only said about each of the classes that the acoustic pictures, belonging to them, are characterized by a certain internal similarity, and at the same time it is possible to distinguish them from the representation of the signals belonging to the other classes. Thus in the discussed case the recognition is connected with detecting and specifying what in fact joins particular individual pictures in specific groups, the number of which, as well as characteristic features, are not known in advance. It is, what should be stressed one more time, a task specifically different from recognising sound picture in the context of semantic identification of a statement or in the context of recognizing the speakers. In the context of this task attempts to apply new procedures methodologies, referring to the concept of automatic **understanding** of pictures, instead of recognising them automatically, are undertaken [18], however, this new and long-reaching important issue will not be discussed in this work broader.

3. Ways of gathering empirical material

All the theses presented in this work are based on particular empirical researches, having been performed, as it has already been mentioned, in the period of several years, on the basis of several various problem areas, separate from the medical point of view, but moving to a such common element, that each time a source of information and the basis of the undertaken decisions are the analysis, based on researching the signal of the pathological speech. In relation to this, one of the important elements, integrating the research described here, is searching the right space structure of the distinctive features for the correct representation of pathological speech signals in all the further discussed tasks of their analysis and recognition. Moreover, solving the stated tasks required at each time gathering appropriate number of samples of these signals, connected with

various forms of considered pathologies. In order to acquire undisturbed results, enabling to get precise and often simply subtle evaluation of quality and usefulness of particular set of entering parameters, it was necessary to have at disposal samples of a very high quality signal. Because of this reason all the acoustic research, on which the suggestions and conclusions presented in this work are based, were performed in a deadroom, registering records was performed by the professional equipment and using solely professional, thoroughly checked, acoustic analysers.

Speech articulation of people was evaluated by a speech test, taking into account these forms of a signal generation and its articulation, which were selected as providing the biggest amount of diagnostic information (as in their generation and articulation the selected, for the purpose of research, fragments of vocal-creating tract take part). The selection of pronounced phrases and sets of words by researched people was performed, based on the morphological and functional analysis of the expected (in the considered pathology) speech organs dysfunction, which caused the fact that in the research there were in disposal the materials including sets of words, selected in such a way, regarding their phonetic features, that they bring the maximum of information. It is worth adding that because of the research specificity (quite often performed on small children with the developmental faults, or on the patients after hard operations in the range of speech organs) it was aimed to provide the least nuisance of the patient's samples recording. Thus the text words were comprehensible, easy to repeat and were composed of one or two syllables.

The final acoustic material has been collected from 175 persons divided into two groups:

- the reference group (the standard group), 25 persons with a correct pronunciation,
- the group of patients (150 persons) treated by the following surgery methods.

Both the patients and the persons from the reference group pronounced the same text (three times), which consisted of: vowels (/a/, /u/, /e/, /i/), words containing vowels (/ala/, /as/, /ula/, /ela/, /igwa/), and a test sentence “/dzis’/ /jest/ /wadna/ /pogoda/” (the weather is nice today).

The product of the initial transformation of the recorded signal was a dynamic spectrum $W(i, j)$, digitized in time, frequency and amplitude by output circuit of the acoustic analyzer. In order to standardise the research process, and in order to provide the results comparability, the same scheme of signal transformation, with the amplitude resolution quant $\Delta s = 1$ dB and with an evenly frequency digitized signal in the waveband, $f_d = 125$ Hz, $f_g = 12$ kHz every $\Delta f = 125$ Hz was applied. It is worth noticing that the frequency division of the spectrum was linear, and not logarithmical, which can be explained by the fact that subtle spectrum deformations, distinguishing a pathological speech from the correct one, are often located in these usually neglected high-frequency areas, treated (rightly) as not very useful in the speech recognition from the semantic or individual point of view. This is one of many particularities distinguishing the considered issues from the pathological speech transformation from the typical issues of the correct speech analysis and recognition. Instant spectra were determined with the usage of a digital analyser and were timely digitised (sampled) every $\Delta t = 9$ ms. The applied

professional registration system provided a transfer band from 20 Hz to 20 kHz at the dynamics amounted to not less than 80 dB.

4. Signal parameterisation

The dynamic spectra $W(i, j)$ achieved as a result of the analysis, were used in the described research directly, as distinctive features vectors in the analysis and evaluation of the pathological speech signal, particularly at the stage of the preliminary research, when there was a necessity to detect the essence of the abnormality of the speech signal time-frequency structure generated by a person suffering from one of the forms of the researched pathologies. However, because of the redundancy and a considerable dimension of such a generated features space, they were usually transformed further, by applying into the research the vectors of features $\mathbf{X}_i$ and of other forms, discussed below.

4.1. Spectrum parameters, disregarding time factor

In a lot of research it was sufficient to bestow attention to single columns of matrix $W(i, j)$, or to the vectors achieved as a result of averaging instant spectra in the determined time intervals. In such an approach a structure of the considered feature vector, being an entry for analysis and identification algorithms was following:

$$\langle f_1, f_2, f_3, \dots, f_{96} \rangle = \mathbf{X}_1, \quad (1)$$

where f_i – averaged amplitude of j dynamic spectrum waveband.

4.2. ‘Traditional’ parameters referring to speech recognition techniques

Despite, repeatedly stressed in his work, analysis methods distinction and pathological speech recognition in relation to the speech analysis and recognition for the semantic purposes or for the biometrical speaker identification, there were attempts of pathological speech signals analysis using the parameters, which confirmed their utility in the traditional analysis methods and in the speech recognition. These are the parameters related to the formants and spectra moments:

$$\langle F_1, F_2, F_3, M_0, M_1, M_2 \rangle = \mathbf{X}_2, \quad (2)$$

where F_1, F_2, F_3 – formant frequencies, M_0, M_1, M_2 – spectra moments introduced and described in previous works.

4.3. Parameters related to signal power in selected bands

One of the observations, which can be made while observing pathology effects in the sphere of speech signal articulation, is a statement that in the pathological signal,

in a considerably different way (in relation to the correct speech pattern), the energy is resolved inside each of the signal spectrum bands. This observation inclined to apply, in some research, a vector of acoustic features, characterised by the following structure:

$$\langle M_0, M_1, M_2, WS_s, WS_1, WS_2, WS_3 \rangle = \mathbf{X}_3, \quad (3)$$

where WS_s – relative power coefficient determining the ratio of signal power in the pattern phoneme band (determined on the basis of statistical research of undeformed speech), to signal power in the whole band of pathological speech signal, WS_i – relative power coefficient determining the ratio of signal power in the i -band ($i = 1, 2, 3$) to signal power in the whole band (selection of the released bands boundaries was one of the main research problems solved within the framework of the research presented here).

4.4. Parameters directed to detect sound generation abnormalities

As well known, a speech signal can be considered as results of a sound generation process (in larynx with a usage of voice folds vibration or above-larynx with applying noise generation into intentionally or casually with created voice route narrowing), and a process of the generated sound modulation using resonance cavities, being created in the speech organs as a result of a jaw, tongue, lips or the soft palate movements. In case of the voice characteristics, described above, the main factor, which was researched and acoustically evaluated, was the modulation process.

Speech signal pathology, however, is very often connected with voice-creation process disturbances (particularly a speech vocal components generation in a pathologically changed larynx), that is why in a number of research the most appropriate one proved to be the features vector, characterized by the structure described below, directed to detect the phonation process abnormality:

$$\langle M_0, M_1, M_2, C_w, C_p, J, S \rangle = \mathbf{X}_4, \quad (4)$$

where C_w – the relative power coefficient, denoting the ratio of signal power in the reference phoneme frequency range to the signal power in the whole frequency band of the signal, C_p – the relative power coefficient, denoting the ratio of the signal power in the remaining frequency band to the signal power in the whole frequency band of the signal, J – **Jitter**, denotes a frequency deviation of the basic tone – in consecutive periods, S – **Shimmer**, denotes an amplitude deviation of the basic tone – in consecutive periods.

4.5. 'Complete' vector

At the end it is worth to mention about a possibility of applying in research the feature vector, which includes the characteristics presented above, and also some additional constituents, enabling to catch the issue of the speech signal process deformation,

independently from where it occurs and what its nature is. Such a ‘complete’ features vector, being used in the presented research, is characterised by the following structure:

$$\langle M_0, M_1, M_2, F_1, F_2, F_3, F_4, AF_1, AF_2, AF_3, AF_4, \\ WS_1, WS_2, WS_3, C_1, C_2, C_3, C_4, C_5, F0_SR, J, S \rangle = \mathbf{X}_5, \quad (5)$$

where AF_1 – AF_4 – formants’ amplitudes values, $F0_SR$ – basic frequency medium value of larynx tone.

The correctness of the space of features selection was verified by means of the cluster analysis.

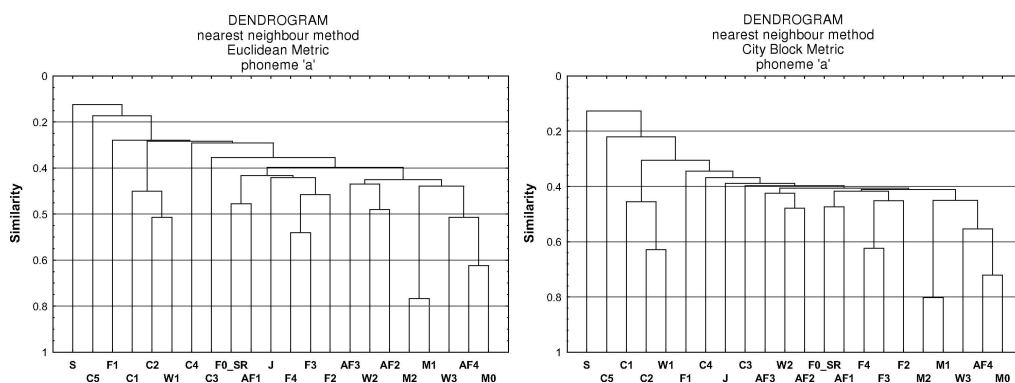
4.6. *Summing up remarks concerning signal parameterisation matter*

The presented above features vectors, used in the reported research, have certainly a diverse range of usefulness. The features, mentioned and characterized above were chosen on the basis of years of research concerning various evaluation measurements of speech deformation degree. In the research, the following features, comprising three following advantages were sought:

- being little sensitive to the statement content and to individual features of the speaker’s voice,
- demonstrating high sensitivity while diversifying various forms of speech pathologies and while classifying various degrees of the same pathology type advancement,
- being easily determined on the basis of the recorded speech signal samples and demonstrating the desired numerical stability (are little sensitive to small errors in the signal measurement).

It seems that the sets, presented above, realize the given task in the highest quality standard. Unique and accurate quality evaluation for the proposed set of distinctive parameters determined for the samples of pathological speech is very difficult, because the phonetic data collected from the examined persons differ also in aspects being outside the scope of our analysis (e.g. the articulation rate is varied), similarly as the statements of persons in good health and correct (standard) articulation are different, regarded as a reference data in the present study. Already at the preliminary stage of the research work it has been found that the quality evaluations of the selected parameters are difficult for quantitative estimation because of the multitude of measured acoustic parameters and also because of the great variety of registered acoustic phenomena. It is also very difficult to construct for various feature vectors a quality measure general enough to be used as a criteria function in the search described here. Exactly this (i.e. the inaccessibility of formal definition of a mathematical quality measure) was the reason that made the application of any automated or computer aided optimisation methods impossible (e.g. gradient search method, but also Monte Carlo methods or genetic algorithms techniques) in the considered task. Therefore in the study described here the evaluation has been done by the authors themselves and the structural analysis technique has been

used instead of quantitative measures. Specifically the cluster analysis has been used as a preliminary method of data analysis, allowing at least a rough estimation of the distribution of objects representing the particular speech samples in the multidimensional space of the analyzed acoustic features. For analyzing the correctness of the feature space structure central agglomeration procedure and seven combinatorial methods have been used. Exemplary dendrograms have been presented in Fig. 1.



was determining the optimal, for the pronunciation, lower front teeth setting in case of toothless patients. Specifically in this research series (the results of which were published separately) the best setting of the lower teeth in the complete dentures was sought for the purpose of the best pronunciation of the 'S' phoneme. The task, which was given to the system of speech pathology degree evaluation, meant in this case the assistance in selecting the most suitable dentures shape, on the basis of the speech signal deformation degree. The exemplary phonemes' courses obtained for the samples of the correct and pathological speech with dentures and without dentures, used in the research are shown in the Fig. 2.

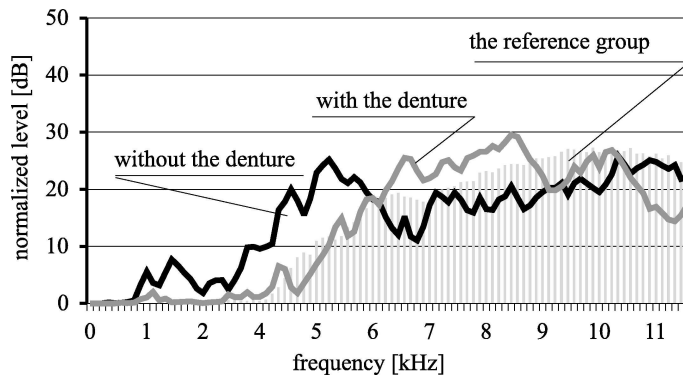


Fig. 2. Typical spectrogram of averaged /s/ phoneme.

5.2. Maxillofacial surgery

Post-surgical losses of a palate, jaw, as well as palatoschisis losses cause a hard aesthetical and functional disability for a patient. The appearance improvement and a considerable decrease of speech functions, chewing, swallowing and breathing sightness is achieved by surgical reconstructions and prosthetic rehabilitation. However, even applying a number of surgical operations and using the most effective prosthetic reconstruction methods, does not always enable to rebuild the lost anatomical structures of a face and mouth completely. It results in the fact that the speech of a patient, after surgical operations and dentures, deviates considerably from the correct speech pattern. Thus there are many indications to apply dentures rehabilitation. Also in this case applying objective (acoustic) evaluation of the speech deformation degree, after various types of operations (quite often medically equivalent, which means giving identical therapeutical effect), can support the process of decision making by a surgeon (certainly in case of the following operations). Speech deformation degree evaluation (and more specifically, evaluation of the limit degree evaluation occurring after the operation of speech deformation) constitutes also a very important element in the denture selection and optimisation. The examples of deformed phonemes courses with marked confidence ranges, for the statement before the surgery, after the surgery, without and with dentures, are presented in the Fig. 3.

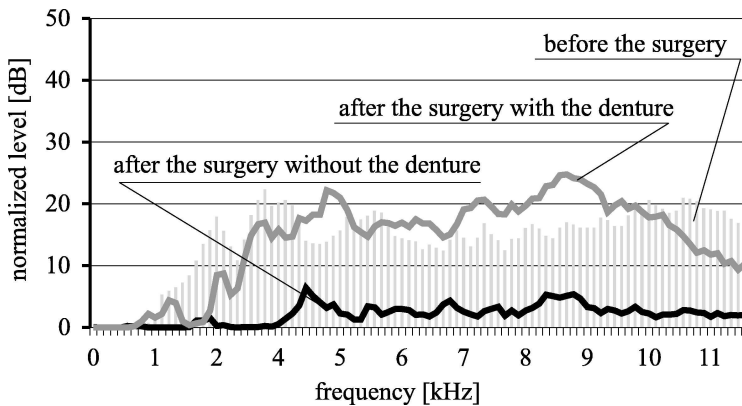


Fig. 3. Averaged spectrogram of /s/ phoneme – before surgery, after surgery without denture, and with denture.

5.3. Laryngology

Because of the often occurring in practise a possibility of selection one of several available surgical methods of larynx medical cure, there is often a necessity of a sudden evaluation of the expected speech deformation degree after the surgery. Undertaking this topic a data concerning forms and speech signal deformation degree, after various surgical operations types was gathered to gain, so-called teaching set, for further applying techniques of artificial intelligence, requiring a learning process (amongst others,

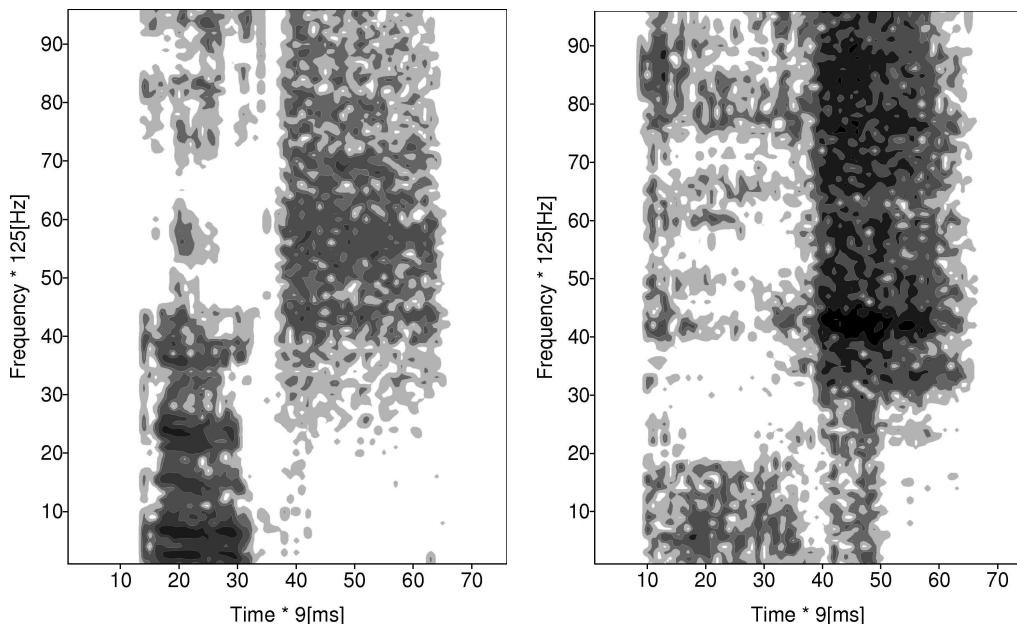


Fig. 4. Spectrogram of /as/ statement for reference group and group of patients.

neuron networks). Speech articulation researches were performed on people suffering from the larynx cancer, men after various types of surgery. Examples of differences in a statement in the reference group and the group of ill people are presented in the spectrograms in the Fig. 4.

In the Fig. 5, examples of a single 'A' vowel spectrum for the reference group and after selected pathologies are presented.

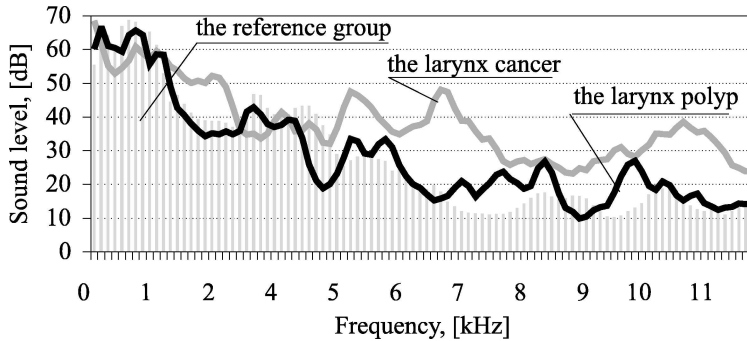


Fig. 5. Spectra of /a/ vowel, correct and deformed by larynx cancer and larynx polyp.

6. Methods of pathological speech signals analysis and classification

The presented above attempts' results for applying automatic grouping methods and concentration analysis algorithms for various (discussed earlier) problems in the range of pathological speech analysis, allow for a certain global orientation in the topology of each of the mentioned above features spaces. Performing such analysis for various sets of signals parameters, constituting a space basis, in which the analysis is performed or in which the recognition is going to be performed, a general orientation concerning these spaces is achieved. In particular, confronting the results of automatic grouping objects in features space with a known (from a medical documentation) systematic of the analysed illness cases, connected with particular signal samples, being the researched form of speech pathology, we can evaluate the selected parameters usefulness and adequacy, and trace (in a synthetic way) the data typology in the adopted features.

The final measure of a particular data and parameters usefulness is, however, always confronting them with the attempt to use them for the final task, which in this considered problem is automatic classification and recognising achieved pathological speech samples. In considering the recognition methods and results defined here, it is worth taking into consideration the following factor: in the considered tasks the process of recognising pathological speech signal acoustic deformation means subordinating the acquired acoustic pictures, resulting from research, to one of the classes, usually unknown before. It is worth to pay attention to this detail and to stress its importance, as in the case of statement recognition, for example in order to steer particular machines or devices by voice comments, the set of recognized classes of pictures is given in advance. Therefore

in the described case the recognition is linked to detection and specification of what essentially joins each individual picture in particular groups, the number of which, and the feature characteristics are not known in advance.

7. Selected research results

The selected results of research are presented below. The degree of speech signal deformation in relation to the dentures parameter to the pattern was presented in the Fig. 6.

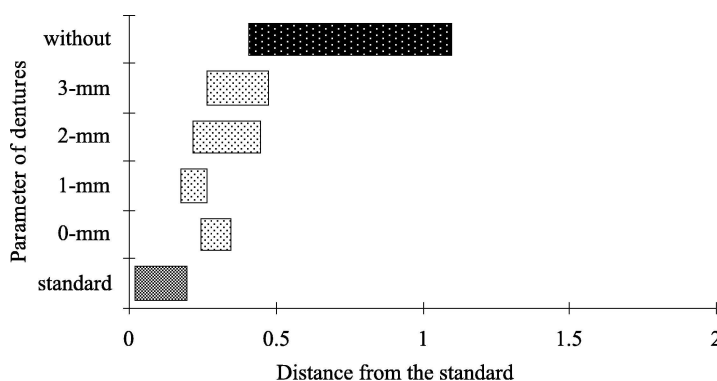


Fig. 6. Distance dispersion between pattern and pathological speech.

In the Fig. 7 it was presented how it is possible to evaluate speech signal deformation metrically, the deformation which was caused by cancer changes in the mouth area and these changes correction by the dentures application, in relation to the correct speech, by applying various metrics.

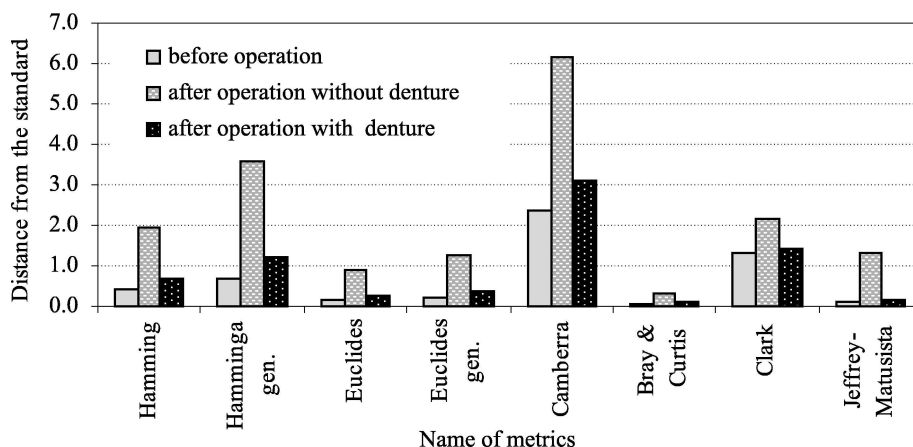


Fig. 7. Distance from examination, before operation and after operation.

In the Fig. 8 the deformation of vowels in relation to the correct speech in case of people suffering from various illness changes in the larynx area is shown [10].

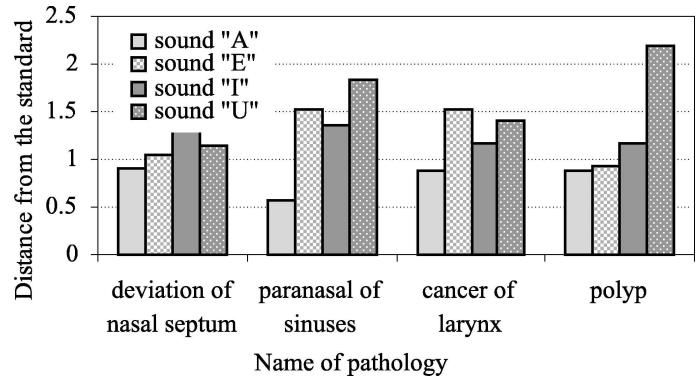


Fig. 8. Distance from standard in case of Cammber’s formula – examination before operation.

In the Fig. 9 there was shown the degree of a vowel deformation in relation to the correct speech in case of a person suffering from larynx cancer in various periods of time, before the surgical operation, after the operation and after a long period of rehabilitation.

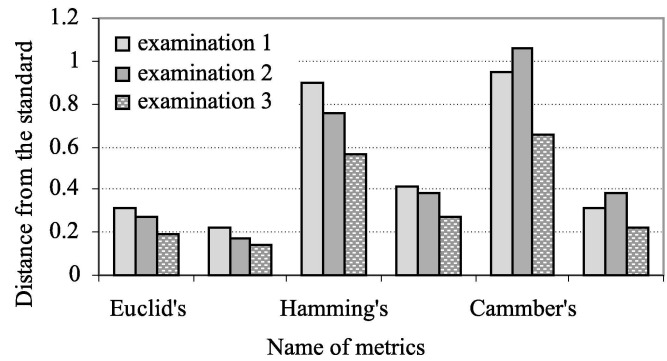


Fig. 9. Distance from standard in case of Cammber’s formula, examination 1 – before operation, examination 2 – after operation, examination 3 – six months after operation.

Figures 6, 7, 8 and 9 “distance vaules from the the standard” are presented as the normalised distance value between the standart speech and patological speech according to established metrics in the chosen space of features.

In the below figures (Fig. 10 and Fig. 11) it is shown how speech is deformed in case of patients suffering from larynx cancer in relation to the kind of performed surgical operation to the correct speech.

The classification and recognition results, achieved on the basis of the pathological speech parameters (vector X_5), taken into consideration in this work, are presented be-

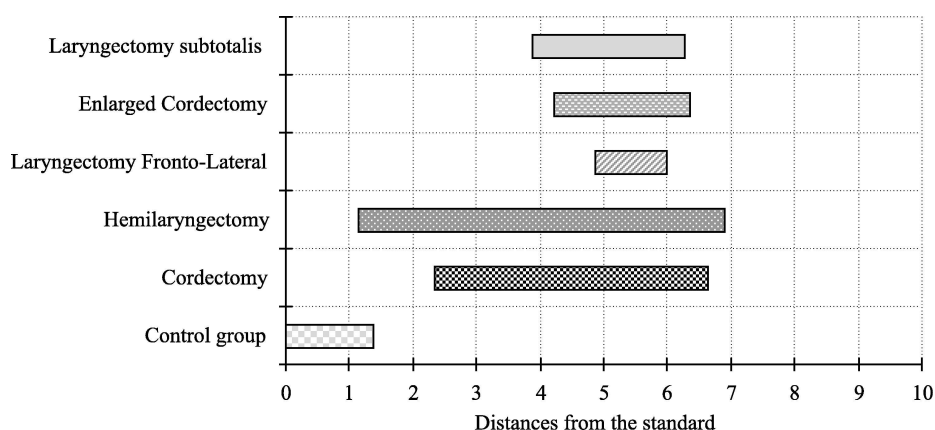


Fig. 10. Degree of dispersion and distribution along the axis marking distance from the standard in case of generalized Hammig's formula.

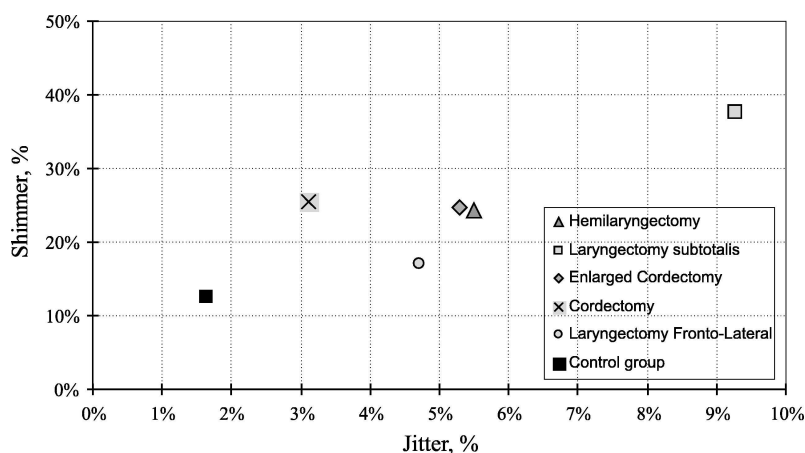


Fig. 11. Surgery of glottis, distribution on the Jitter–Shimmer scale.

low. To achieve a higher degree of reliability the qualification and recognition tests were performed by applying two methods: pattern recognition algorithms, originating from the artificial intelligence, and by applying neural networks, based on the biocybernetics achievements [6].

The presented results definitely confirm the opinion that the selected parameters characterise well the analysed phenomena connected with the speech signal degradation forms in the context of selected pathologies.

In the Table 2 there are presented the results of research in the form of correct recognitions (expressed in %) for the neural networks, applied in four groups of research materials, connected to laryngology [14].

The presented results can be recognised as satisfactory, what proves that the selected parameters generally well characterise the analysed pathological speech signal deformation phenomena.

Table 1. Results obtained by various pattern recognition methods, Euclidean Metric (Prosthodontia – 35 patients, Maxillofacial surgery – 24 patients).

Recognition method	Recognition reliability [%]	
	Prosthodontia	Maxillofacial surgery
NN algorithm	84	90
α -NN algorithm	76	90
NM algorithm	84	85
Optimal spherical neighbourhoods	–	79
Quadratic approximation	67	82

Table 2. Results obtained by neural networks methods (hemilaryngectomy and enleng. corpectomy – 28 patients, laryngectomy sub. – 23 patients, corpectomy – 14 patients, laryngectomy fronto-lateral – 10 patients).

No of neurons in the input layer	Hemilaryngectomy and Enleng. Corpectomy [%]	Lar. Subtotalis [%]	Corpectomy [%]	Lar. Front. Lat. [%]
96	73	72	70	67
6	79	77	75	74
7	86	86	84	83

8. Conclusions

By analysing the research results obtained in individual groups (research materials), our task is not to prove that this or that structural parameter of the dentures is better or worse, but to prove that the acoustic method, together with the metrics, enables us to obtain, from the various patient's statements (with various structural parameters of denture), the particular one, which we will consider the most similar to the pattern. It is possible to evaluate the signal deformation degree, caused by illness changes (before the surgical operation), and the deformation degree after the surgical operation (e.g. big loses in the mouth area) and the degree of deformation improvement caused by having applied dentures.

In the case of an operation on a section of the larynx, based on the results of tests discussed previously, we can decide on the type of operation (hemilaryngectomy (Fig. 11) – enlarged corpectomy). We take currently an investigation on the influence of the intubations on the voice quality [19]. The goal of the investigations is the extraction of the patient' group with a higher risk of the vocal cords damages as a result of the intubations.

The presented methodology can be directly applied in the monitoring examinations for patients after larynx surgery as well as the operations of the nose and paranasal sinussis and intubations in general anaesthesia.

The results of the present study seem to be helpful in proper qualification of patients for the specific types of operations, in order to provide at the same time the maximal therapeutic effect and minimal speech deformation degree after the operation.

Appearing, as a consequence of the performed research, a usefulness of the features sets (signal description parameters), dedicated especially to the task of evaluating the pathological speech, confirms the thesis, that to this particular task one should use the dedicated features, and not the commonly used parameters and signal description elements, used in the correct speech signal analysis and recognition. The revealed, at this opportunity, a dominancy of neuronal networks in solving tasks formulated here, is a quite obvious consequence of the generally known suitability of that tool in the tasks characterised by a big form complexity of areas subordinated to particular classes in the features. This result indicates a possibility to apply the neuronal networks technique as a profitable alternative to other techniques (pictures recognition or statistical methods), but, it is not the main subject matter of the consideration. However, it is worth to stress, one more time, the key role of the correct preparing the parameters, describing the researched speech signal, as in the case when it is impossible to parameterise a problem in the way, which would guarantee a good resolution in the features of the areas subordinated to individual classes, the results, achieved by optional methods (especially the techniques of pictures recognition) are highly unsatisfactory.

It is worth to add that despite the basic character, orientated to analyse the properties of various features elements, the research, performed by us, are helpful in the practical matters, in the dentures and brace selection, and also in qualifying the patients to particular types of surgical operations, so that it is possible that they deform the operated patient's voice in a small degree.

Acknowledgments

The author would like to thank the Ministry of Science and Higher Education, Poland for financial support (grant 4 T11E 031 23).

References

- [1] ADAMCZYK J. A., WSZOŁEK W., *The analysis and ewaluation of machine emitted vibroacoustics signal patterns*, Zeszyty Naukowe AGH, Mechanika, **9**, 4, 5–15 (1990).
- [2] DEMENKO G., *Analiza cech suprasegmentalnych języka polskiego na potrzeby technologii mowy*, UAM, Poznań 2000.
- [3] ENGEL Z., MODRZEJEWSKI M., WSZOŁEK W., *Akustyczna ocena operacji krtani z wykorzystaniem parametrów tonu podstawowego*, Zeszyty Naukowe AGH, Mechanika, **16**, 1, 1997.
- [4] GROCHOLEWSKI S., *Statystyczne podstawy systemu ARM dla języka polskiego*, WPP, Poznań 2001.
- [5] GUBRYNOWICZ R., *Komputerowe modelowanie artykulacji głosek języka polskiego*, IPPT PAN, Warszawa 2000.
- [6] IZWORSKI A., TADEUSIEWICZ R., WSZOŁEK W., *Artificial intelligence methods in diagnostics of the pathological speech signals*, [in:] NEGOITA M. GH., HOWLETT R. J., JAIN L. C. [Eds.],

- Knowledge-Based Intelligent Information and Engineering Systems*, Lecture Notes in Computer Science, vol. 3215, pp. 740–748, Springer-Verlag, Berlin – Heidelberg – New York 2004.
- [7] MAKAREWICZ R., *Wstęp do akustyki teoretycznej*, Wyd. UAM, Poznań 2005.
- [8] MAJEWSKI W., MYŚLECKI W., BRACHMAŃSKI S., *Metody oceny jakości transmisji sygnału mowy*, OSA 2000, pp. 61–72, Rzeszów – Jawor, 19–22.09.2000.
- [9] MODRZEJEWSKI M., OLSZEWSKI E., WSZOLEK W., REROŃ E., STREK P., *Acoustic assessment of voice signal deformation after partial surgery of the larynx*, Auris Nasus Larynx, International Journal of ORL & NNS, Japan, **26**, 183–190 (1999).
- [10] MODRZEJEWSKI M., OLSZEWSKI E., STREK P., WSZOLEK W., ZIELIŃSKA J., *Effectiveness of classical chordectomy in the treatment of cancer of the glottis*, Auris Nasus Larynx, International Journal of ORL & NNS, Japan, **25**, 59–66 (1998).
- [11] OZIMEK E., *Dźwięk i jego percepcja. Aspekty fizyczne i psychoakustyczne*, Wydawnictwo Naukowe PWN, Warszawa – Poznań 2002.
- [12] REROŃ E., TADEUSIEWICZ R., MODRZEJEWSKI M., WSZOLEK W., *Application of neural networks and pattern recognition methods to the evaluation of speech deformation degree for patients surgically treated for larynx cancer*, Neuroendocrinology Letters, **19**, 3, 147–157, 1998, Mattes-Heidelberg-Germany.
- [13] SZALENIEC J., MODRZEJEWSKI M., WSZOLEK W., *Research on the influence of endotracheal intubation on the speech signal*, Speech Analysis Synthesis and Recognition in Technology Linguistics and Medicine, AGH, , Krakow 2005, pp. 127–133
- [14] TADEUSIEWICZ R., IZWORSKI A., WSZOLEK W., WSZOLEK T., *Processing and classification of deformed speech using neural networks*, Proceedings of The First Joint Meeting of BMES & EMBS Serving Humanity, Advancing Technology, 13–16 Oct., Atlanta, USA 1999.
- [15] TADEUSIEWICZ R., IZWORSKI A., WSZOLEK W., *Pathological speech evaluation using the artificial intelligence methods*, World Congress on Medical Physics and Biomedical Engineering, September 14–19, Nice, France 1997.
- [16] WSZOLEK W., *Pewne miary stosowane w ocenie mowy zniekształconej*, Mat. Konf. I Krajowa Konferencja Głosowa Komunikacja Człowiek–Komputer, pp. 201–206, Wrocław 1985.
- [17] WSZOLEK W., TADEUSIEWICZ R., *Methods of voice signal analysis using artificial intelligence methods*, Archives of Acoustics, **28**, 3, 245 (2003).
- [18] WSZOLEK W., WSZOLEK T., *Automatic understanding of speech signal*, [in:] KŁOPOTEK A., WIERZCHOŃ S., TROJANOWSKI K. [Eds.], *Intelligent Information Processing and Web Mining*, pp. 609–620, Springer-Verlag, Berlin – Heidelberg – New York 2004.
- [19] WSZOLEK W., MODRZEJEWSKI M., *Methods of acoustic voice signal evaluation after ENT surgery*, 1st International Conference on Experiments/Process/System Modelling/Simulation/Optimization, Athens, 6–9 July, 2005 (on CD).
- [20] WSZOLEK W., KŁACZYŃSKI M., *Acoustic methods of voice estimation after surgical treatment of the vocal tract*, Archives of Acoustics, **30**, 4 (Supplement), 193–197 (2005).