ARCHIVES OF ACOUSTICS
21, 1, 3–13 (1996)

USEFULNESS OF THE ACOUSTIC SPEECH SIGNAL INFORMATION IN THE NERVOUS SYSTEM EVALUATION

CZ. BASZTURA* AND R. PODEMSKI**

* Institute of Telecommunication and Acoustics
Wrocław Technical University I-28
(50-370 Wrocław, Wyspiańskiego 27)
** Department and Clinic of Neurology
Wrocław Medical University

In the paper a preliminary analysis of the possibilities of using a speech signal for the objective diagnostics assistance, as well as for monitoring the process of the reconvalescence from the nervous system diseases, was presented. This paper presents also a conception of the acoustic speech signal computer analysis for some nervous pathological states. The possibilities of making acoustic patterns for the dysarthria disease states from distinctive acoustic parameters compared to the physiological standards, are also described in the papers. The paper presents conditions required for making speech pathology that can determine the kind, and in some cases, even the localization of the nervous system changes. The examples of the acoustic pictures of different types of dysartrhia compared to the psychological standards, were also described in the papers. The examples of the acoustic patterns of different types of dysartrhia compared to the psychological standards were described, and future research directions are set.

1. Introduction

Human speech is a set of the specific sound signals that result from a transformation taking place on different levels:

a) intellectual, connected with the process of thinking and initial verbalization of the information perceived reversibly and creating the so called inner speech [25],

b) semantic, where the verbal elements (words, phonemes) acquire certain meaning,

c) linguistic, connected with the phonetical and grammatical form of a statement,

d) articulatory, connected with the process of making and forming of speech sounds characteristic of a given language and conditioned by the correct function of a certain nervous system, speech and structures of the breathing organs [21, 23, 26],

e) acoustic, i.e. the physical representation of speech.

Signals of speech sound, set in order according to the rules of the given language, are the conveyors of the substance that we want to transmit the complex process of communication with other people. They include personal information (e.g. development level, sex) and are certain reflections of the emotional conditon as well as complex functions of the central and peripheral nervous systems and the speech organ of the speaker. All these elements, along with the transformations mentioned above, have an influence on and are built in the speech signal's acoustic structure [4, 24]. Thus, except for the information connected directly with the communicative function (the substance of the statement), the speech sound signals include nonverbal elements seemingly redundant or useless in the process of verbal communication. However, a number of these elements is closely connected with the pronounced substance giving an appropriate emotional tone to it and to the sound characteristic of the given language (the prosody of speech) [20, 25]. To sum up, it may be stated that the acoustic signal of speech includes two essential information categories:

a) the verbal-semantic category that determines substance of the statement and makes the ground for communication by means of speech (It has to be observed that this category refers not only to the acoustic speech signal but also to the "soundless" written or gesture speech),

b) the nonverbal categories, that penetrates — so to say — all the elements of the pronounced speech and determines the physical structure of speech as an acoustic signal.

The first of the given above categories of information (a) is fundamental for the process of communication between men or between man and computer, while the other one (b) is a database which, after performing a computer acoustic analysis, can be used e.g. for voice identification and verification of the speech disturbances diagnostics [19, 20]. The range of instantaneous value changes of the pressure level of the sound signal of speech is of tens of decibels, the frequency ranges to 10 000 Hz, which in effect gives an information efficiency estimated to about 250 000 bits per second.

2. Basic acoustic properties of the speech signal

Speech being a complex acoustic phenomenon is the result of functions of particular central and peripheral nervous systems, articulation muscles and the breathing system [21]. Individual perception of the phenomena contributing to the picture of speech is impossible by means of the sens of hearing. The application of a precise computer analysis with graphic and digital description makes it possible to acquire many data that are not available by direct auditory perception connected with the function of anatomical structures initiating the speech process, stimulating it and giving the speech sounds their ultimate acoustic shape [5, 6, 7, 9, 13, 19, 20]. The possibility of exploration of these structures by means of conventional research methods is often limited and is mostly based on a personal view of the researcher.

In order to utilize the speech signal for objective diagnostic assistance as well as for monitoring the course of convalescence from diseases related to the nerve system one has to take into consideration the definitions of particular pathological states and characteristic recurrent features of the speech signal of people in good health (physiological acoustic standard of speech signal). The computer analysis of the acoustic signal in the defined pathological states will make it possible to distinguish the features typical of a particular syndrome by comparison with the physiological standard [7, 8, 11, 12, 22]. Then the conditions are made for the creation of acoustic speech pathology models which would indicate the kind and, in some cases, even the localization of the nerve system's changes [19, 20].

The task to determine a connection between pathological states and certain acoustic symptoms included in the speech signal that have a diagnostic significance is very complex because, among others, of the fact that the described transformations that take place on different levels of the nervous system (cortex, sub-cortex structures, brain stem, nerves, neuro-muscular connections i.e. synapses, articulation muscles) [1, 2, 14, 21]. This requires the introduction of an acoustic comparative survey on every stage of the signal analysis. This survey may be based, in the initial phase of research, on spectrograms, time parameters as the course of envelope or energy, the density of zero crossings etc. [19, 20]. The essential method task before the beginning of an acoustic speech research is to determine the optimal set of "key entries" i.e. the forms of speech (test tasks) that would deliver information concerning the changes within the widely recognized speech organ (central and peripheral elements of the nervous system, executive structures — voice apparatus), showing at the same time the least possible sensibility to other features (as the personal or social-linguistic ones etc.). The starting point for the sound speech signal characteristics is the thesis that perceptible acoustic features are primarily connected to the voice descriptions that take into consideration the following elements:

1) the elementary pitch period of the vocal cords vibration tone,
2) laryngeal tone peak amplitudes,
3) the frequency spectrum ranges of the speech signal,
4) the frequency ranges of noise components in the speech signal,
5) the amplitude envelope decrease or increase,
6) the time of duration (abbreviation, elongation) of the speech sequence.

Objective acoustic diagnostic methods are now widely applied in phoniatrics. They are based on the speech signal information value analysis as natural and ultimate product of the speech organ operation [3, 10, 15, 16]. These methods support conventional phoniatric studies. Sometimes they show their advantages as they:

a) are varried out in normal physiological conditions of phonation and articulation,

b) do not need any introduction of auxiliary instruments or foreign matters to the patient's body and therefore are not traumatizing,

c) enable the real time visualization of chosen speech signal acoustic parameters on a TV monitor or any other peripheral device of a computer system,

d) the results of the research can be recorded, e.g. as a computer printout. This enables an objective comparative estimation of the course of the therapy disease and the course efficiency.

The functions of the speech organ, being a matter of interest for the phoniatrists, are determined by stimulation and cooperation of some structures in the nervous system. The elimination of the executive speech organ primary disesase e.g. inflammatory or tumorous process within larynx as well as the determination of physiological and pathological models of the acoustic speech signal makes it possible to apply the computer speech analysis in noninvasive, sensitive and objective diagnostics of nervous system function disturbances [8, 11].

3. Acoustic analysis of speech disturbances of nervous origin

Cortex systems controlling the speech process are located in the dominating (mostly the left one) cerebral half. Impairement or handicap of these systems functions leads to a complex speech disturbance called aphasia which is the subject of interdisciplinary neurological, neuropsychological, linguistic research [18]. Dysarthria is the other crucial category of speech disturbances. Its acoustic expression is a speech signal's deformation taking — in extreme cases — the form of anarthria that makes it impossible do identify speech sounds [17]. Dysarthria results from an incorrect transfer of impulses from the cortex of the dominating cerebral half to the articulation muscles through nerve pathways (central part), nerves and neuro-muscular synapses (peripheral part). It is also the result of sufficiency of the articulation muscles as well as of the systems modulating the movement functions connected with the speech process cerebellar and extrapyramidal systems [21, 23]. Dysarthria of the cortex origin is a very rare phenomenon. Selective damage of certain movement areas of the dominating cerebral half is the reason for this speech disturbance. The individual types of dysarthria have a series of characteristic features that enable their identification by means of audition:

1) cerebral dysarthria — faint, "neglectful" articulation, the pace of prononunciation gives it a scanning nature. The damage concerns cerebellum or nerve pathways connecting the cerebellum with other structures of the nervous system [1],

2) bulbar dysarthria — speech is "blurred", with a distinct nasal sound. The damage may concern central nervous paths transmitting impulses from the movement cortex of both cerebral halves to the medulla oblongata, i.e. bulbous, concentrations of nerve cells (nuclei) in the bulbous, nerves supplying the articulation muscles, neuro-muscular synapses and the muscles [14],

3) extrapyramidal dysarthria — the speech is low, "shivering" often accelerated and faint (in Parkinson's disease) or less frequently it has "burtsting" character (in Huntington's disease). The damage concerns under-cortex nuclei and/or the extrapyramidal system's nerve pathways,

4) dysarthria in myasthenia — in the course of speaking the voice lowers, articulation aggravates as in bulbous dysarthria. This state is originated by weakness

of the articulation muscles caused by the impulse transmission from nerve to muscle capability exhaustion within synapse.

Considering the acoustic phenomena appearing in different types of dysarthria, this kind of speech disturbances seems to be richest source of diagnostic information. It is possible to obtain these information by a properly programmed computer speech signal analysis.

The initial stage of research includes the preparation of test tasks taking into consideration the widest possible range of articulation schemes having a suitable, for the given language, difficulty level (noise, voiced, fricative, plosive sounds put together into sequences that realization requires full articulation proficiency). During the research carried out at the Department and Clinic of Neurology of the Medical University of Wrocław and the Institute of Telecommunication and Acoustics of the Wrocław Technical University a basic sentence: "The maintenance technician tightens radiator" (in Polish: "Konserwator uszczelnia kaloryfer" — [*konservator uʃtʃelnŋa kalorƎferƎ]*) was used [19, 20]. It was a kind of a key entry which, after recording it on a magnetic tape, was subject to computer analysis. The construction of the test sequence, apart from articulation possibilities of a man, should take into account the diagnostic purpose that one wants to achieve.

The spectrograms of a test statement in case of the cerebellar dysarthria show characteristic phenomenon of "jitter" demanding further analyses (Fig. 1). The acoustic pattern of bulbar dysarthria is characterized by "disorganisation" of the sound spectrum; selectivity of each element of a statement deteriorates, the energy of the voice diminishes, the time to perform subsequent stages of the test task elongates. An increase of the formants having higher frequency can also be observed (Fig. 2).

In the extrapyramidal speech disturbances the main element is the loss of amplitude (decrease of voice energy), particularly clearly visible when compared to the test pronouncement of a healthy man (Fig. 3).

The sentence which is captious in the sense of articulation "Konserwator uszczelnia kaloryfer" ("The maintenance technician tightens radiator"), can cause difficulties to the people with cerebellar or bulbar dysarthria, while for those with myasthenia, where it is essential to discover or evidence of muscle weakness, the oral test should include repeating (therefore provoking the fatigue of certain group of muscles) articulation structures. A series of pronounced digits may be a good example of it "jeden i jeden" ("one and one"), "jeden i dwa" ("one and two"), "jeden i trzy" ("one and three"), "jeden i cztery" ("one and four") etc. Owing to miasthenic weakness of muscles, the "tired" articulation structure refuses to work properly. This in turn leads to disorganisation of the speech signal which, in the acoustic picture of speech, is characterized by increasing abbreviation of the unit pronouncement time and loss of amplitude. After intravenous injection of a drug causing a fast but brief improvement of the neuro-muscular conductivity (Tensilon), one can register renewed gain of amplitude, the change of the structure of particular phonemes and elongation of the test speech pronouncement time taking the form similar to that one of healthy people (Fig. 4) [19, 20].

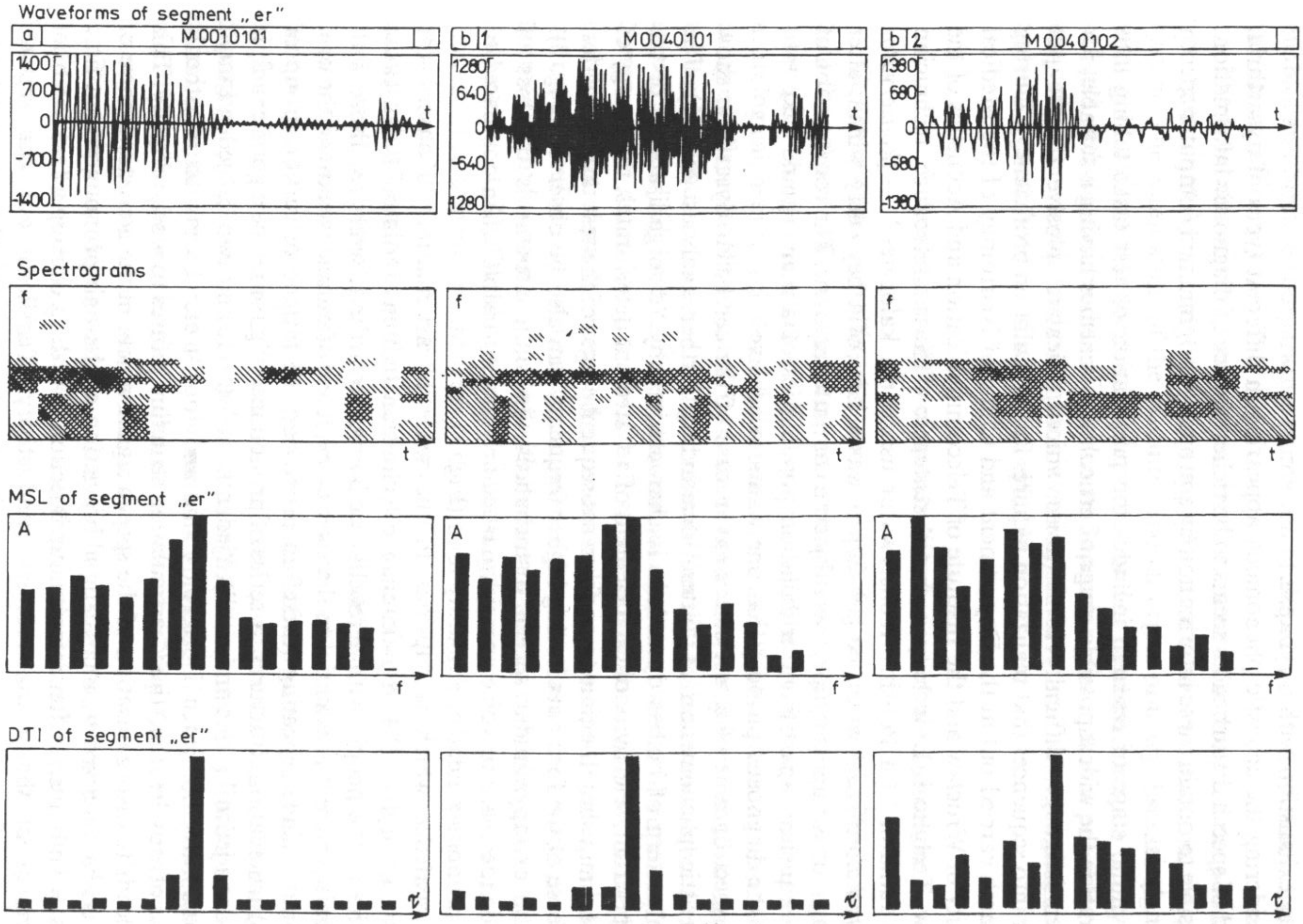

Fig. 1. The waveforms, spectrograms, mean spectrum levels (MSL), and distribution of time intervals between zero-crossings of the speech signal (DTI) of diade "er" of "konserwator", spoken by male voice; a — normal voice, b1 — patient with middle form of cerebellar dysarthria, b2 — patient with severe form of cerebellar dysarthria.

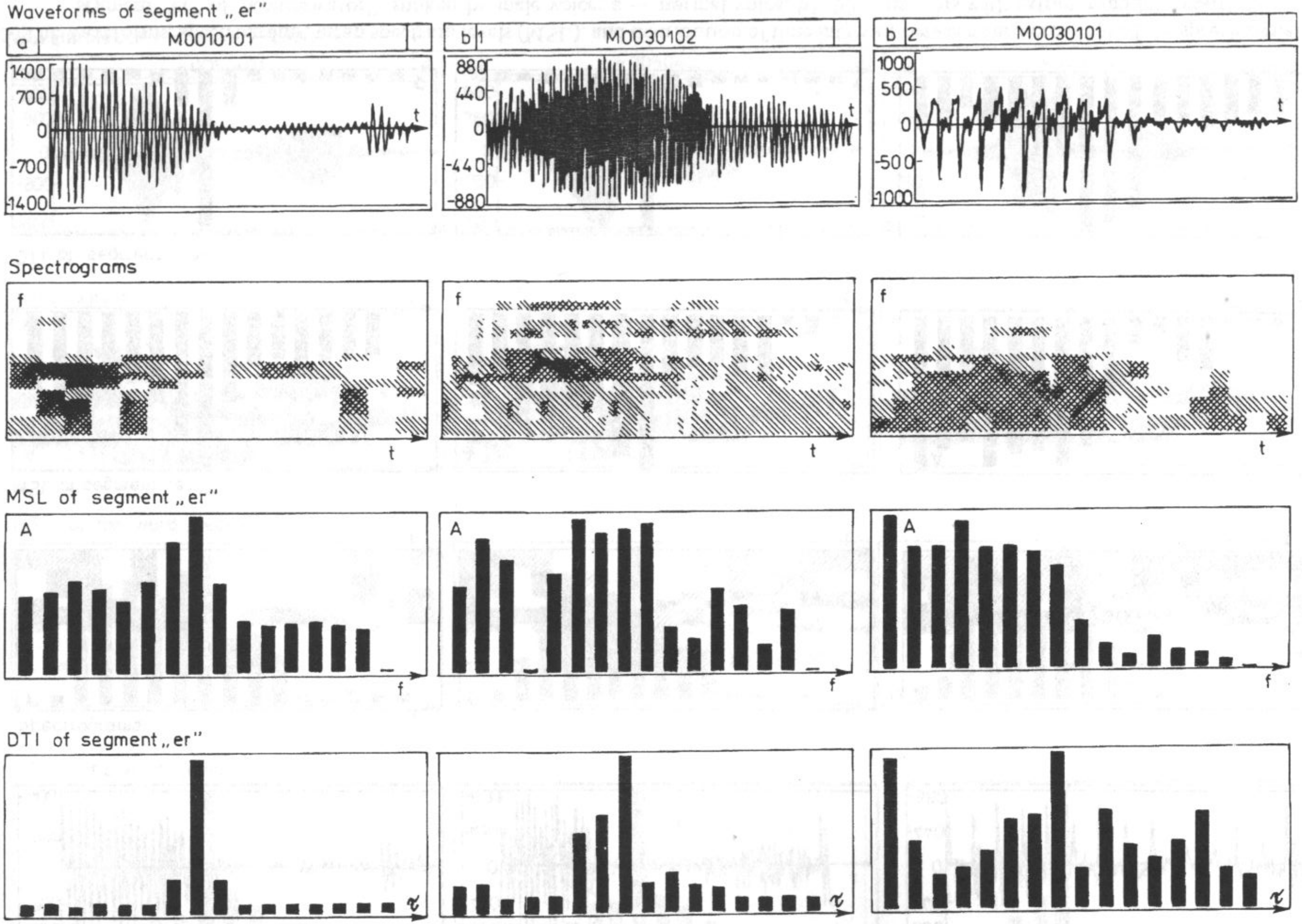

Fig. 2. The waveforms, spectrograms, mean spectrum levels (MSL), and distribution of time intervals between zero-crossings of the speech signal (DTI) of segment "er" of "konserwator", spoken by male voice; a — normal voice, b1 — patient with middle form of bulbar dysarthria, b2 — patient with severe form of bulbar dysarthria.

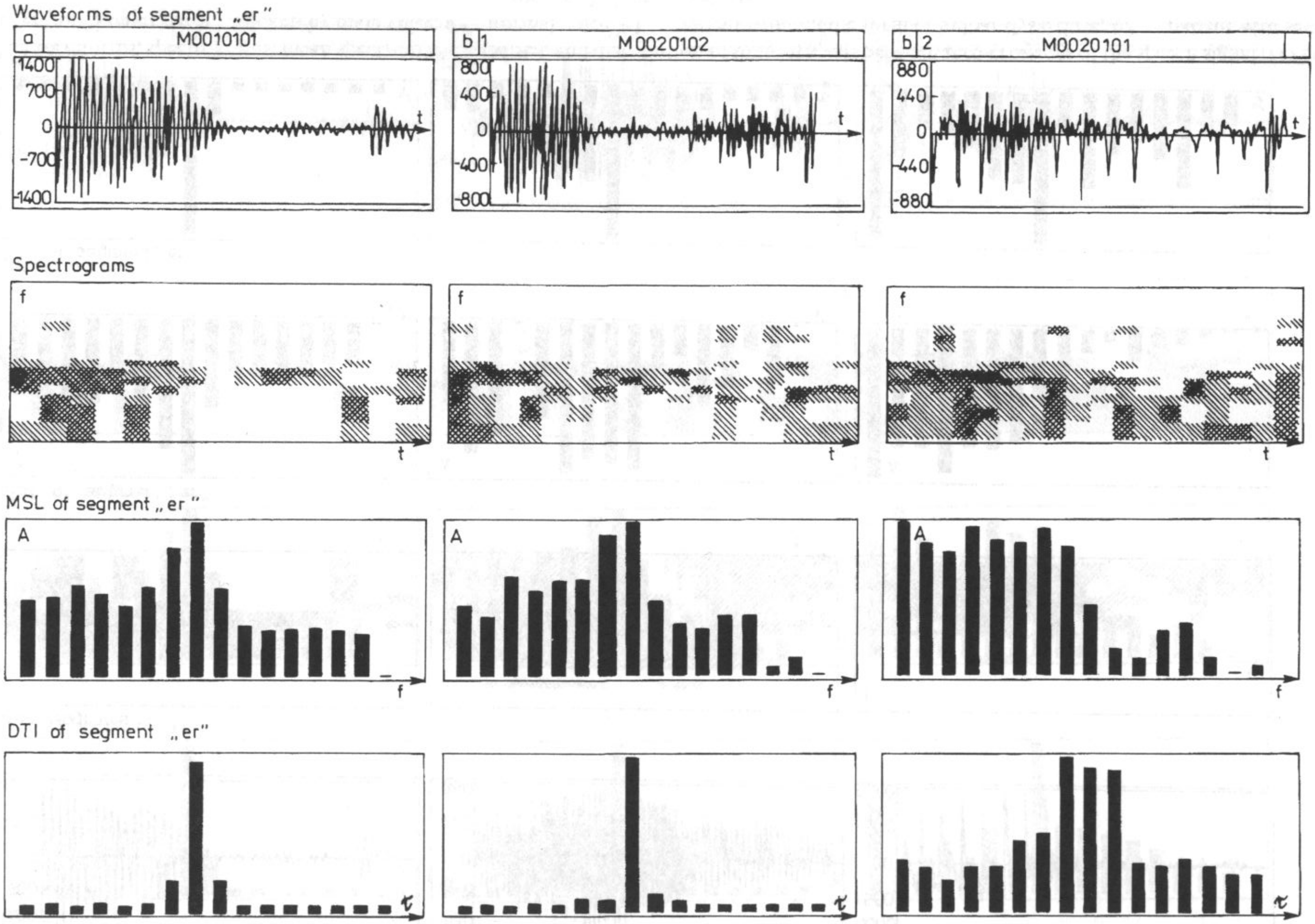

Fig. 3. The waveforms, spectrograms, mean spectrum levels (MSL), and distribution of time intervals between zero-crossings of the speech signal (DTI) of segment "er" of "konserwator", spoken by male voice; a — normal voice, b1, b2— patients with extrapyramidal dysarthria.

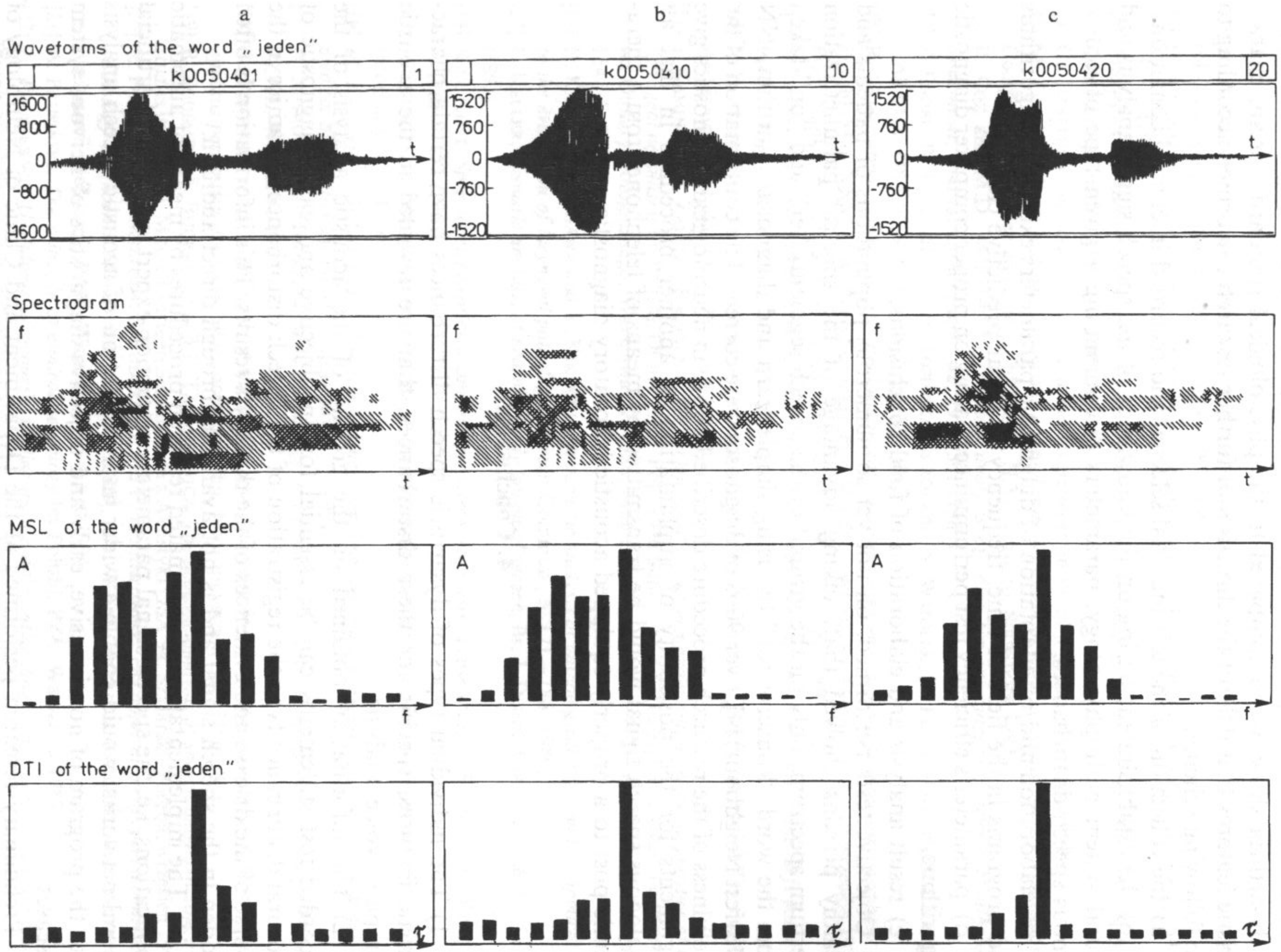

Fig. 4. The waveforms, spectrograms, mean spectrum levels (MSL), and distribution of time intervals between zero-crossings of the speech signal (DTI) of word "jeden" [*jeden*]; a, b, c — patient (woman) with dysarthria in myasthenia — successive 1-th, 10-th and 20-th repetitions.

The described examples of the speech acoustic patterns evidencing different types of dysarthria show, when compared to the physiological standard, certain characteristic features that delimit the direction of further research performed according to the following scheme:

a) the collection of the test material selected and recorded in clinical conditions,

b) the establishment basing on the knowledge of the speech signal analysis and neural system pathophysiology, parameters discriminating a given type of neurogenous speech disturbances,

c) elaboration and/or adaptation of diagnostic parameters extraction algorithms and programs in the field of time, frequency or linear predictive coding,

d) parameter's efficiency test performances basing on proper computer diagnostic algorithms,

e) result analysis and elaboration of final conclusions.

Objective tests performed, on not yet so numerous populations of patients and healthy persons showed that, taking advantage of the spectral parametrization (tertium spectrum) only on the ground of the speech segments "er" and "on" taken from the word "konserwator" by using simple learn and diagnosis algorithms NN (Nearest Neighbour) [4], over 90% of diagnoses were correct. The confirmation of the usefulness of the research procedure described above in neurological diagnostics give arguments for the possibility of implantation of another procedure in that an analogous speech signal would be transmitted by means of telephone from a therapeutic post to a properly equipped acoustic laboratory diagnostic center.

4. Conclusions

1) The individual types of dysarthric speech disturbances have certain characteristic features, specific for these disturbances, that are revealed in the acoustic computer voice analysis.

2) The information obtained on the ground of the acoustic analysis of the recorded test statements can be essential for an objective and early diagnostic of a neural disturbance type the registration of the speech disturbance dynamics in the course of the disease and evidences of the therapy's results. This information is often hidden in the speech signal and is not available through direct audition.

3) The implementation of automated research procedures by means of automatic applications, i.e. the speech signal parameters extraction expert systems and artificial neural networks, would enable a wider implementation of acoustic speech analysis into the program of non-invasive, early and objective diagnostics of nervous system diseases.

Preliminary research performed at the Department and Clinic of Neurology of Medical University of Wrocław and at the Institute of Telecommunication and Acoustics of Wrocław Technical University give every reason to believe that this program is perfectly real.

References

[1] H. Ackermann, *Articulatory deficits in parkinsonian dysarthria: an acoustic analysis*, J. Neurol.-Neurosurg.-Psychiatry, **54**, 12, 1093–1098 (1991).

[2] H. Ackermann, *Cerebellar voice tremor: an acoustic analysis*, J. Neurol.-Neurosurg.-Psychiatry, **54**, 1, 74–76 (1991).

[3] Alku-Paavo, *Glotal wave analysis with synchronous iterative adaptive inverse filtering*, Speech Communication, **11**, 109–118, (1992).

[4] Cz. Basztura, *Sources, signals and acoustics patterns* [in Polish], WKL, Warszawa 1988.

[5] M. Cooke, S. Beet and M. Crawford, *Visual representation of speech signals*, Willey Professional Computing, J. Willey and Sons, Chichester, N.Y., Toronto, Singapore 1993.

[6] B. Cox and M.D. Morrison, *Acoustical analysis of voice for computerized laryngeal pathology assessment*, Journal Otolaryngol., **12**, 295–301 (1983).

[7] Z. Engel, R. Tadeusiewicz, H. Tosińska-Okrój and W. Wszołek, [in Polish], Otwarte Seminarium z Akustyki OSA-92, Kraków 1992, pp. 95–98.

[8] R. Gubrynowicz, *The acoustic analysis of the substitutive speech in patients after laryngectomy* [in Polish], Mat. XXVI Otwartego Seminarium z Akustyki, Wrocław–Oleśnica, 17–19 wrzesień 1979, Prac. Nauk. Inst. Telek. Pol. Wrocławskiej 43, Ser. Konf. 11, pp. 189–192.

[9] J. Kacprowski, W. Mikiel and A. Szewczyk, *The acoustic model investigations of the cleft palat* [in Polish], Arch. Akust., **11**, 2, 167–187 (1976).

[10] J. Kacprowski, *The objective acoustic methods in the diagnostics of organ of speech* [in Polish], Archiwum Akustyki, **14**, 4, 287–304 (1979).

[11] J. Kamiński, M. Zaleska-Krecicka, *The technique of the speech examination in the diagnostics of the laryngeal neoplasms* [in Polish], Otolaryngol. Pol., 1986 Supl. — XXXIII Zjazd Polskiego Towarzystwa Otolaryngologicznego, Warszawa 29, 1986, Materiały Naukowe, v. 1, Leczenie raka krtani, pp. 164–1.

[12] H. Kasuya, Sh. Ogawa and Y. Kikuchi, *An acoustic analysis of pathological voice and its application to the evaluation of laryngeal pathology*, Speech Commun., **5**, 171–186 (1986).

[13] E. Keller, *Acoustic analysis of neurologically impaired speech*, Br. J. Distrd. Commun., **26**, 1, 75–94 (1991).

[14] R.D. Kent, *Speech deterioration in amyotropic lateral sclerotsis; a case study*, J. Speech Hear. Res., **34**, 6, 1269–1275 (1991).

[15] J. Kulikowski, *The computer analysis of the pictures in biology and medical diagnostics* [in Polish], Informatyka, 1, 1–4 (1991).

[16] M. Kurzyński, J. Sas and E. Puchała, *Unconventional methods of recognition in the computer aid medical diagnostics* [in Polish], Informatyka, 7, 1991.

[17] A. Mitronowicz-Modrzejewska, *Physiology and pathology of voice, hearing and speech* [in Polish], PZWL, Warszawa 1963.

[18] R. Podemski *et al.*, *Character of the middle and late components of the acoustic evoced potenti in aphasus perception disorders*, Electrocephal. Clin. Neurophysical., **52**, 21P (1981).

[19] R. Podemski, S. Budrewicz and P. Tarantowicz, *The acoustic analysis of speech diagnostic the pilot studies* [in Polish], Mat. Otwartego Seminarium z Akustyki OSA'94, Szklarska Poręba, wrzesień 1994.

[20] R. Podemski, Cz. Basztura, S. Budrewicz and P. Tarantowicz, *The application of speech analysis in pathology diagnostics of nerves system* [in Polish], Mat. Konf. Naukowa Komitetu Biocybernetyki i Inżynierii Medycznej PAN, Warszawa 12–13.10.1994.

[21] A. Prusiński, *The principle of clinical neurology* [in Polish], PZŁ, Warszawa 1989.

[22] A. Pruszewicz, *Usefulness of acoustic studies on the differential diagnostics organic and functional dysphonia*, Acta Otolaryngol. Stockholm, **111**, 2, 414–419 (1991).

[23] R. Tadeusiewicz, *The speech signal* [in Polish], WKiL, Warszawa 1988.

[24] L.S. Wygodski, *The thinking and speech* [in Polish], PWN, Warszawa 1981.

[25] A. Zakrzewski, *The clinical otholaryngology* [in Polish], PZWL, Warszawa 1982.