ARCHIVES OF ACOUSTICS
4, 4, 289-304 (1979)

OBJECTIVE ACOUSTICAL METHODS IN PHONIATRIC DIAGNOSTICS OF SPEECH ORGAN DISORDERS*

JANUSZ KACPROWSKI

Institute of Fundamental Technological Research,
Polish Academy of Sciences (00-049 Warszawa, ul. Świętokrzyska 21)

Physical parameters of the speech signal contain essential information on the anatomical structure and the kinematics of the biological sound source, i.e. the human speech organ. Several objective acoustical methods have been proposed and developed recently, which not only assist, but in some cases are even superior to the classical, mostly intrusive and/or subjective diagnostic methods, commonly used at present in clinical practice in laryngology and phoniatrics.

The subject of the present paper is the theoretical fundamentals and the engineering performance of some selected methods and systems for speech signal processing and analysis, aimed at acoustical diagnosis of larynx and vocal tract pathology. The current status of fundamental research in this domain is briefly reviewed and discussed, including the scientific activity and achievements of the Department of Cybernetic Acoustics, IFTR–PAS. Preliminary results of experimental research are given. Special attention is paid to the role of model investigations in the development of acoustical diagnostic methods and systems in laryngology and phoniatrics. The references play a role of the reader's guide throughout the rather scarce literature on the subject considered.

1. Introduction

For years physicians have used two traditional methods for the examination of speech organ disorders and pathology. The first method relies on the subjective auditory evaluation of the qualities of the patient's voice. The second method consists in visual observation of the articulatory effectors using more or less complex techniques, such as direct or indirect laryngoscopy, laryngeal stroboscopy etc. The principal handicap of the auditory methods is their subjectivity and the lack of absolute quantitative standards, although many attempts have been made to qualify and to quantify the subjective

* This paper was written under the nodal research problem 10.4 "The nervous system and biocybernetic systems and elements".

measurements of voice quality defects (cf. PERKINS [29]). The main shortcoming of the visual methods is that they are as a rule uncomfortable, and sometimes even painful to the patient, being besides ineffective for early detection of pathology. Consequently, their application is usually confined to diagnosis during the later stages of the disease and to surgical situations. More advanced techniques, such as high-speed motion picture analysis, glottography, electromyography, pneumatachography etc. have mostly research applications.

New objective methods in phoniatric diagnostics became feasible in the early sixties of this century, when FANT [4] presented his acoustical theory of speech production and FLANAGAN [8] summarized the recent results of speech communication research in his book on speech analysis, synthesis and perception. Since then the speech signal has been considered – from a physical point of view – as the output of the concatenation of three independent acoustical filters, which represent the glottal source, the vocal tract and the mouth and/or nose radiation impedance, respectively. Moreover, the advanced computer technique has offered new conspicuous facilities in speech signal processing and analysis. Consequently, the research work in the new field of applied acoustics, viz. the acoustical diagnostics of speech organ disorders, was initiated and has been henceforth intensively developed, offering many possibilities of practical applications not only in the medicine: laryngology and phoniatrics, but also in acoustic phonetics, applied linguistics and even in vocalistics.

2. General characteristics of acoustical methods

Acoustical diagnostic methods in laryngology and phoniatrics, which are based – generally speaking – on an analysis of the information content of the speech signal as the final and natural output of the speech organ, have recently been applied in clinical practice [30-35]. They not only support, but in some cases are even superior to the conventional methods, since

– they are used under normal conditions of phonation and articulation;

– they neither need any surgical tool nor foreign substance to be introduced into the patient's body, nor have to be aided by any painful and dangerous intrusive procedure;

– they enable the real-time visualization of the acoustic parameters of the speech signal on the TV monitor, computer display or other conventional peripheral device of a computer system, and are thus especially convenient during rehabilitation process, since the loss of hearing, usually associated with speech disorders, may be compensated by the additional and auxiliary information channel of the sight organ;

– the printed output of the computer-aided analysis of the speech signal may be inserted into the patient's medical history and used afterwards for the objective evaluation of the effectiveness of the therapy.

Acoustical methods consist in the measurement and analysis of those physical parameters of the speech signal which are tightly correlated with the anatomical structure and with the kinematics of the biological source, i.e. the human speech organ. The development and application of these methods in clinical practice must be based on the understanding of the acoustics of speech production and supported by the knowledge of versatile and universal physical models of the biological system under consideration, i.e. the larynx source and the vocal tract. In fact, modelling offers the best means of getting an insight into the functional relations between the acoustical parameters of the speech signal and the intrinsic structural and kinematic features of the speech organ under both normal and pathological conditions.

The directly measurable diagnostic signal is the acoustic pressure $p(t)$ of the speech wave at a definite point on the symmetry axis of the subject's mouth. This signal may be considered, at least during the articulation of voiced speech sounds with glottal excitation, as the response of a linear and passive transmission system which represents the vocal tract and is described by its transfer function $H(s)$, being submitted to the action of the impulse excitation function of the larynx source, i.e. the volume velocity $U_g(s)$ of air in the glottis and loaded by the radiation impedance $Z_r(s)$. These relations are schematically shown in Fig. 1.

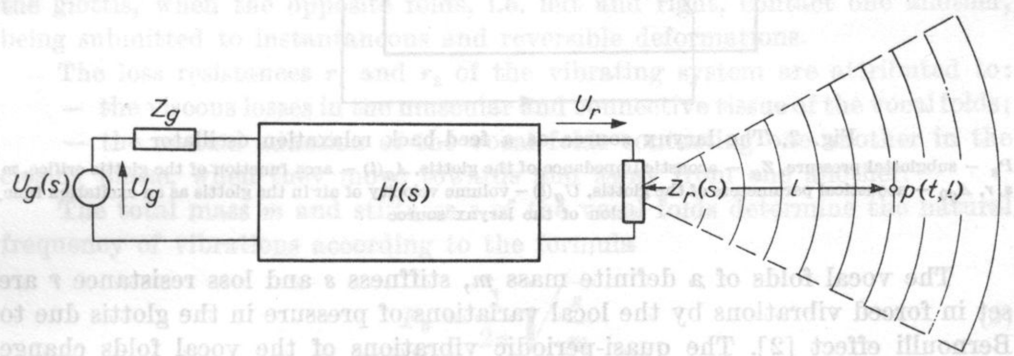

Fig. 1. Equivalent electrical circuit of the speech organ during the articulation of voiced speech sounds with glottal excitation

$U_g(s)$ – excitation function of the larynx source, Z_g – internal impedance of the larynx source, U_g – volume velocity of air in the glottis, U_r – volume velocity of air in the mouth orifice, $H(s)$ – transfer function of the vocal tract, $Z_r(s)$ – radiation impedance of the mouth orifice, $p(t, l)$ – acoustic pressure of the speech wave in the measurement point at a distance l on the symmetry axis of the mouth

The well known general expression for the sound pressure $P(s)$ of the speech wave in Laplace transform presentation according to [4] is given by the equation

$$P(s) = \mathscr{L}\{p(t)\} = U_g(s)H(s)Z_r(s), \tag{1}$$

where Z_r stands for the radiation impedance of the mouth outlet orifice, usually approximated as a circular vibrating piston set in a spherical or infinite flat baffle which simulates the human head [8].

It is evident from (1) that the speech wave reflects the intrinsic features of both the larynx source and the vocal tract. This fact is very promising from the point of view of further diagnostic applications, since the laryngeal pathology and the anomalies of the vocal tract constitute the majority of speech organ disorders most frequently occurring in clinical practice.

3. Physical models of the larynx source

According to the myoelastic-aerodynamic theory formulated by van den Berg [1], the human larynx is – from a physical point of view – a relaxation feed-back oscillator whose output, i.e. the flow or volume velocity $U_g(t)$ of air in the glottis, affects the internal parameters of the system, including those which determine the glottis impedance Z_g. These relations are illustrated schematically in the block diagram shown in Fig. 2.

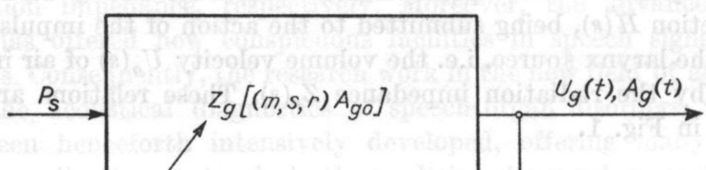

Fig. 2. The larynx source as a feed-back relaxation oscillator

P_s – subglottal pressure, Z_g – acoustic impedance of the glottis, $A_g(t)$ – area function of the glottis orifice, m, s, r, A_{g0} – acoustical parameters of the glottis, $U_g(t)$ – volume velocity of air in the glottis as an excitation function of the larynx source

The vocal folds of a definite mass m, stiffness s and loss resistance r are set in forced vibrations by the local variations of pressure in the glottis due to Bernoulli effect [2]. The quasi-periodic vibrations of the vocal folds change the acoustic impedance Z_g of the glottis and thus modulate the air flow $U_g(t)$ which at the entrance to the vocal tract has the form of recurrent discrete triangle shaped pulses. The latter play the role of an impulse source function $U_g(t)$ which excites vibrations of air in the vocal tract at frequencies corresponding to the resonances of the vocal-tract cavities.

The mechanical vibrating system of the vocal folds may be modelled in many ways [18], starting with the simplest single – mass model with lumped constants and one degree of freedom elaborated by Flanagan and Landgraf [10], up to complex models with distributed constants composed of elementary masses, stiffnesses and loss resistances, as proposed by Ishizaka [15]. Such a sophisticated model is, however, neither necessary nor even convenient in medical diagnostics, since the main physiological features of the human larynx may be simulated with sufficient accuracy in the two-mass model with lumped constants, proposed by ISHIZAKA and MATSUIDARA [16], presented by FLANAGAN [9] to the 7th ICA Congress in Budapest and described in details by ISHIZAKA and FLANAGAN [17].

In the latter model each vocal fold, left and right, is divided into two partial masses m_1 and m_2 which represent its lower and upper parts, respectively. The masses m_1 and m_2 are mutually coupled by the spring s_c which represents the bending stiffness of the vocal folds in the vertical plane parallel to the direction of vibrations. In this way the two-mass model simulates an essential physiological feature of the biological system,. i.e. the deformability of the vocal folds and the phase difference between the displacements of their lower and upper edges in quasi-harmonic motion. The ratio at which the total mass of each vocal fold is divided into its partial components m_1 and m_2, may vary widely depending on the realistic physiological or pathological conditions. The springs s_1 and s_2 associated with the masses m_1 and m_2 represent the individual non-linear stiffnesses of both parts of each vocal fold, i.e. the lower m_1 and the upper m_2 which are opposed to:

– the momentary displacements of the masses m_1 and m_2 from their neutral positions;

– the viscoelastic deformations of the vocal folds in the closing phase of the glottis, when the opposite folds, i.e. left and right, contact one another, being submitted to instantaneous and reversible deformations.

The loss resistances r_1 and r_2 of the vibrating system are attributed to:

– the viscous losses in the muscular and connective tissue of the vocal folds;

– the mutual adhesion of the vocal folds contacting one another in the closing phase, when they move towards and away from one another.

The total mass m and stiffness s of the vocal folds determine the natural frequency of vibrations according to the formula

$$F_0 = \frac{1}{2\pi}\sqrt{\frac{s}{m}}. \tag{2}$$

Under real conditions of phonation and articulation, due to the contraction of the vocalis muscles, the vocal folds are shortened or elongated and their tension Q varies to a limited extent. Their mass m and stiffness s also vary and the pitch or larynx tone frequency changes from the value F_0 given by (2) to the value F_0^1 expressed as follows:

$$F_0^1 = \frac{1}{2\pi}\sqrt{\frac{sQ}{m/Q}} = QF_0. \tag{3}$$

The tension coefficient Q is commonly used in model investigations as an additional parameter which determines the real physiopathological behaviour of the subject's larynx source.

It is evident from this brief description that the two-mass model of the human larynx is versatile enough to take into account and to reflect not only in qualitative, but also in quantitative terms, several laryngeal disorders and anomalies most frequently occurring in clinical practice. The following examples may be quoted. Benign growths, such as vocal nodules and polyps located on one or both folds, increase their mass, whereas the partial atrophy of the vocal fold is equivalent to the reduction of its effective mass. The immobility of one or both vocal folds due to unilateral or bilateral recurrent laryngeal palsy (paralysis) may be expressed in terms of their stiffness value which approaches infinity in spastic cases and reduces to zero in paretic cases.

4. Objective evaluation of laryngeal pathology

It is evident from (1) that the acoustic measures of laryngeal pathology, based on the analysis of the speech wave form $p(t)$, are affected by the vocal tract structure $H(s)$ which, especially at an early stage of pathological development, may distort or even mask some of the important acoustical attributes of larynx disorder. The technique of inverse filtering of speech may be used to remove the influence of the supraglottal structure upon the speech wave form $p(t)$ and to obtain the residue signal whose characteristics are correlated with the movements of the vocal folds only, as it was proposed by MILLER [26]. Thus, if the speech signal contains acoustical information indicative of laryngeal pathology, the measures based on the residue signal should be as informative as the original speech signal itself, if not more so. Valuable information concerning the application of inverse filtering of speech for detecting and evaluating laryngeal pathology may be found in the works of DAVIS [3], KOIKE and MARKEL [22], and MARKEL and GRAY [23].

If the speech signal expressed by (1) is submitted to pre-emphasis equivalent to the reciprocal of the mouth radiation factor $Z_r(s)$ and if an inverse filter is used whose reciprocal is an estimate of the vocal tract transfer function $H(s)$, the result of inverse filtering is the residue signal which is associated with the larynx source excitation function $U_g(s)$ only.

For normal voice, the residue signal has the form of a series of sharp spikes at the initiation of the successive pitch periods T_0, with an irregular or noisy behaviour of relatively low amplitudes between the spikes. The cyclic oscillations within the pitch periods of the primary speech signal have been effectively removed by inverse filtering. The noisy component in the residue signal decays so that during the other half of each pitch period it is much smaller than the spikes at the pitch period initiation. The most important feature of the residue signal for normal voices is thus a high peak signal to noise ratio within the other half of each pitch period.

In pathological cases, on the other hand, especially when an incomplete glottal closure occurs, the assumption of vocal tract and larynx source separability in the linear model of speech production is not valid. As a rule, the more severe the pathology is, the less distinct pattern of periodic spikes can be observed in the residue signal. In an extremely severe case of vocal fold fixation, the acoustic wave form contains a small, if any, periodic component. In the residue signal the cyclic behaviour has been completely eliminated, leaving a very noisy nonperiodic signal. The residue signals of a normal and two pathological voices are shown in Fig. 3, after KOIKE and MARKEL [22], and MARKEL and GRAY [23].

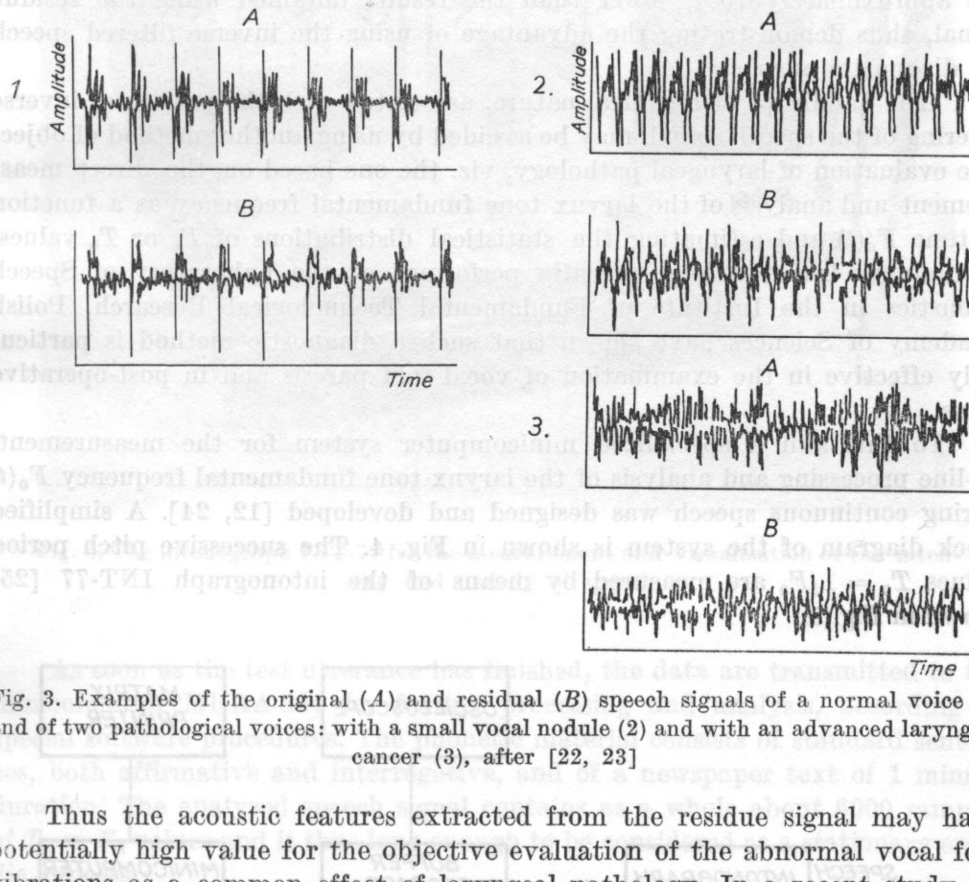

Fig. 3. Examples of the original (*A*) and residual (*B*) speech signals of a normal voice (1) and of two pathological voices: with a small vocal nodule (2) and with an advanced laryngeal cancer (3), after [22, 23]

Thus the acoustic features extracted from the residue signal may have potentially high value for the objective evaluation of the abnormal vocal fold vibrations as a common effect of laryngeal pathology. In a recent study by DAVIS [3], the six features which describe the acoustical structure of the speech signal $p(t)$ previously submitted to inverse filtering, are the following:

— the pitch amplitude of the residue signal, PA;
— the pitch period perturbation quotient, PPQ;
— the pitch amplitude perturbation quotient, APQ;
— the coefficient of excess, EX;
— the spectral flatness of the inverse filter, SFF;
— the spectral flatness of the residue signal, SFR.

These measures were then separately and jointly analyzed for their effectiveness in discriminating between normal and pathological voices. In a closed maximum likelihood test using sustained vowel sounds from 17 normal and 21 pathological speakers, each single feature successfully discriminates between 65 % and 85 % of normal and pathological sounds. The six features jointly produce a 95 % probability of detection for the pathological sounds and a 6 % probability of false alarm for the normal sounds. An analogous joint discrimination test using the unprocessed vowel sounds gave results which are approximately 10 % lower than the results obtained using the residue signal, thus demonstrating the advantage of using the inverse filtered speech for diagnostic purposes.

The handicaps of technical nature, associated with the process of inverse filtering of the speech signal, may be avoided by using another method of objective evaluation of laryngeal pathology, viz. the one based on the direct measurement and analysis of the larynx tone fundamental frequency as a function of time $F_0(t)$ and estimating the statistical distributions of F_0 or T_0 values. Preliminary investigations, recently performed at the Laboratory of Speech Acoustics in the Institute of Fundamental Technological Research, Polish Academy of Sciences have shown that such a diagnostic method is particularly effective in the examination of vocal fold paresis and in post-operative cases.

To this aim a specialized minicomputer system for the measurement, on-line processing and analysis of the larynx tone fundamental frequency $F_0(t)$ during continuous speech was designed and developed [12, 24]. A simplified block diagram of the system is shown in Fig. 4. The successive pitch period values $T_0 = 1/F_0$ are measured by means of the intonograph INT-77 [25] shown in Fig. 5.

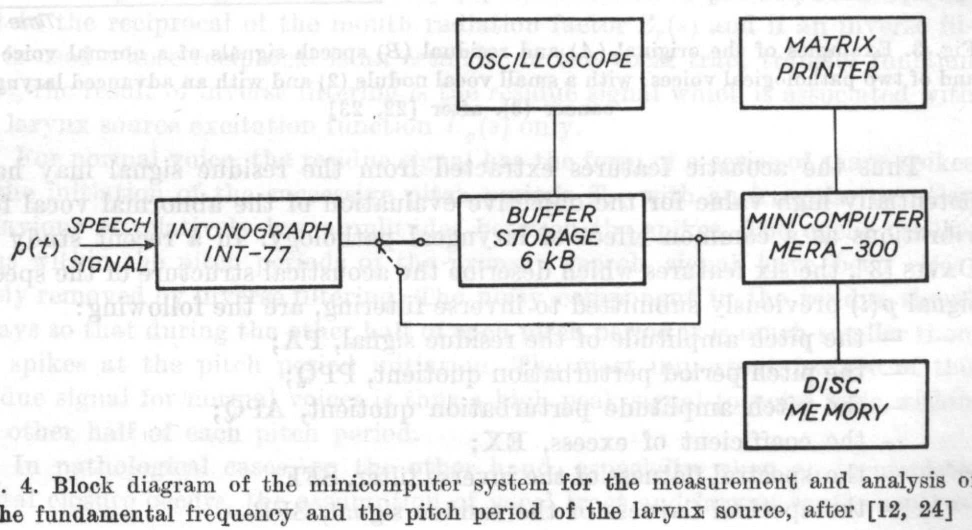

Fig. 4. Block diagram of the minicomputer system for the measurement and analysis of the fundamental frequency and the pitch period of the larynx source, after [12, 24]

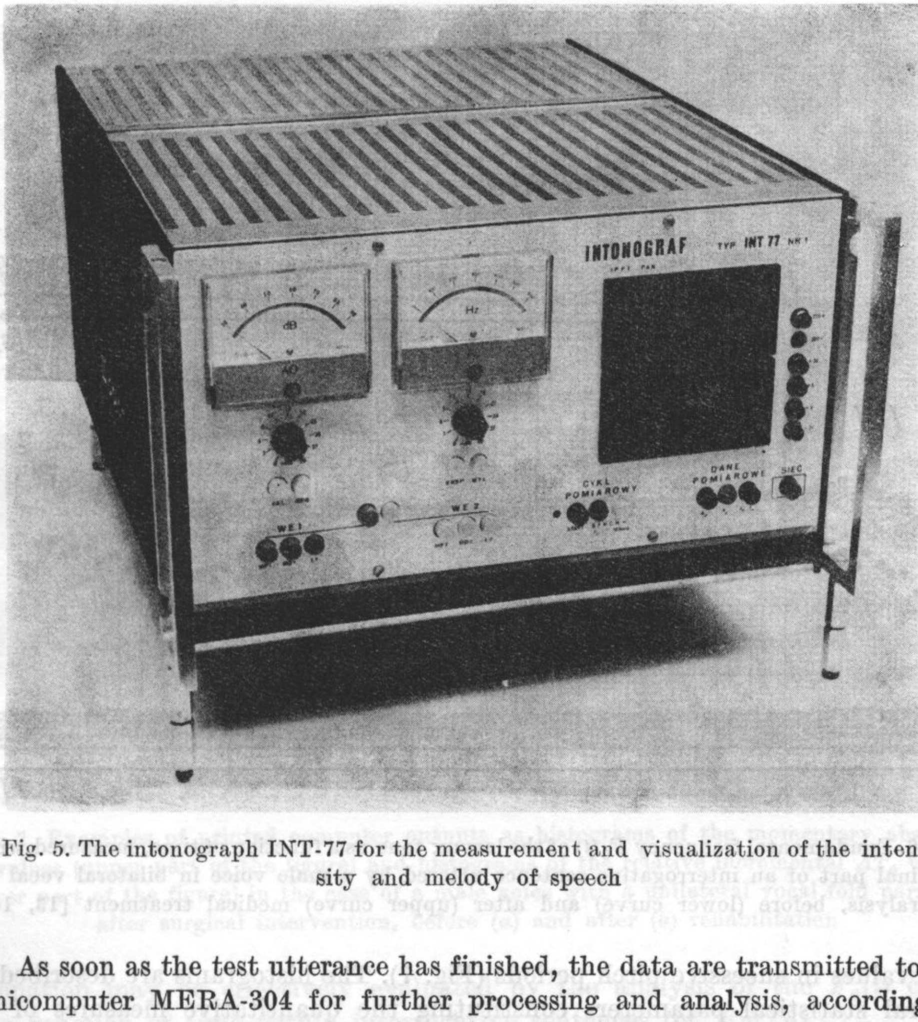

Fig. 5. The intonograph INT-77 for the measurement and visualization of the intensity and melody of speech

As soon as the test utterance has finished, the data are transmitted to the minicomputer MERA-304 for further processing and analysis, according to special software procedures. The phonetic material consists of standard sentences, both affirmative and interrogative, and of a newspaper text of 1 minute duration. The analyzed speech signal contains as a whole about 6000 samples of T_0 or F_0 values and is thus long enough to be considered as a stationary ergodic process.

Two kinds of criteria for the objective evaluation of laryngeal pathology have been accepted. The first measure is the irregularity of the intonation curve $F_0(t)$, which appears chiefly within the final parts of interrogative sentences (Fig. 6). The second measure is the statistical distribution of the T_0 or F_0 values, which is computed by the long-term analysis of the speech signal corresponding to the newspaper text. The results of the statistical analysis are presented as printed output of the minicomputer in the form of histograms of the momentary T_0 values and histograms of the momentary incremental

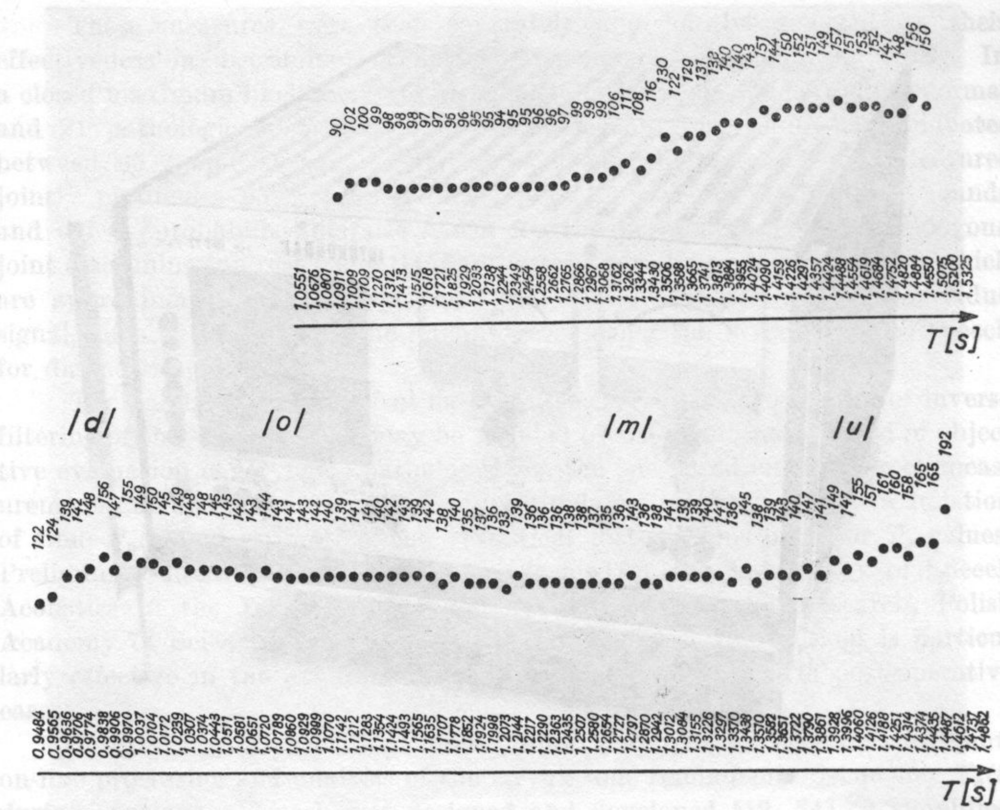

Fig. 6. Fundamental frequency $F_0(t)$ of the larynx tone as a function of time, measured within the final part of an interrogative sentence uttered by a male voice in bilateral vocal fold paralysis, before (lower curve) and after (upper curve) medical treatment [13, 14]

ΔT_0 values in sucessive pitch periods (Fig. 7). The histograms are described by several statistical parameters constituting the quantitative measures of the laryngeal pathology, such as e.g.: mean value $\overline{F}_0$ or $\overline{T}_0$, variance D, standard deviation σ and central moments of the 3rd and 4th order which are used for computing the asymmetry AS and coefficient of excess EX of the T_0 or F_0 distributions [13, 14].

The method was verified under clinical conditions in the Otolaryngological Clinic of the Institute of Surgery, Medical Academy, in Warsaw, on 140 pathological voices with documented larynx disorders. It has been proved that:

— for diagnostic purposes the ΔT_0 distributions are more informative generally, than the T_0 distributions;

— the most effective measures for discriminating between normal and pathological voices are: the asymmetry AS and excess EX of both T_0 and ΔT_0 distributions, as well as the standard deviation σ of the ΔT_0 distribution;

— the inability of a pathological larynx to convey speech melody and

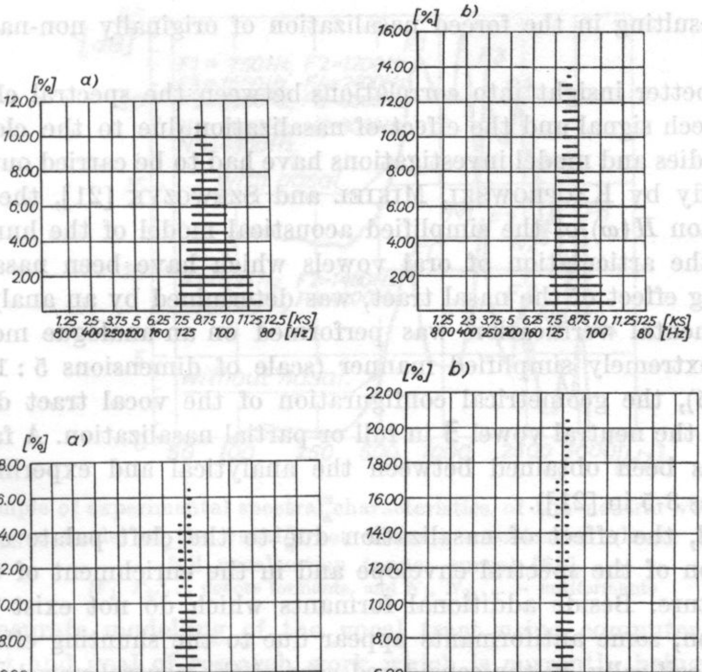

Fig. 7. Examples of printed computer outputs as histograms of the momentary absolute T_0 values (upper part of the figure) and histograms of the relative incremental ΔT_0 values (lower part of the figure) in the case of a male voice with a unilateral vocal fold paralysis after surgical intervention, before (*a*) and after (*b*) rehabilitation

intonation may be effectively evaluated by the analysis of the $F_0(t)$ curve as a function of time within interrogative sentences (cf. Fig. 6).

The results obtained previously are promising and encourage further research. The method and system proposed will hopefully provide facilities for objective computer-aided evaluation of laryngeal pathology in routine clinical practice before and after medical treatement or surgical intervention and during the process of rehabilitation.

5. Objective evaluation of vocal tract pathology

The most frequently occurring pathology of the vocal tract is the cleft palate* (med.: palatoschisis) whose direct effect consists in the permanent and uncontrolled acoustical coupling between the pharynx-mouth cavity and the

* Cleft palate occurs – depending on demographic conditions – in one case for several hundred to one thousand births.

nasal tract, resulting in the forced nasalization of originally non-nasal speech sounds.

To get a better insight into correlations between the spectral characteristics of the speech signal and the effect of nasalization due to the cleft palate, theoretical studies and model investigations have had to be carried out first. In an earlier study by KACPROWSKI, MIKIEL and SZEWCZYK [21], the resultant transfer function $H(\omega)$ of the simplified acoustical model of the human vocal tract during the articulation of oral vowels which have been nasalized due to the shunting effect of the nasal tract, was determined by an analytical method. Experimental verification was performed on an analogue model simulating in an extremely simplified manner (scale of dimensions 5 : 1, scale of frequency 1 : 5), the geometrical configuration of the vocal tract during the articulation of the neutral vowel ∃ in full or partial nasalization. A fairly good conformity has been obtained between the analytical and experimental results (cf. Tables 3-5 in [21]).

In general, the effect of nasalization due to the cleft palate consists in the deformation of the spectral envelope and in the enrichment of the vowel formant structure. Beside additional formants which do not exist in purely oral articulation, some antiformants appear due to the shunting effect of the nasal tract. Specifically, in the considered case of the neutral vowel ∃, the following effects may be observed:

— increase of the first formant frequency by about 12 %;
— additional formants at 1250 Hz and 1700 Hz;
— antiformants at 1400 Hz, 2800-2900 Hz and 4200 Hz;
— disappearance of the second formant $F2$ at a frequency $F_2 = 1500$ Hz due to the appearance of an antiformant at 1400 Hz.

These effects are exemplified by spectral characteristics shown in Fig. 8. The intensity of cleft palate, expressed in terms of a gradual increase of the oral-nasal coupling factor from about 12 % up to its maximal value, affects both the formant-antiformant frequencies and amplitudes, as may be seen in Figs. 10-14 of the previously cited work [21].

The results of our preliminary experiments have confirmed the potential possibility of evaluating the cleft palate by spectral analysis of voiced speech sounds, when using the phonospectroscopic method [30]. The future clinical application of this metod must be based, however, on a better knowledge of a universal articulatory model of the vocal tract, which could be adapted for:

— reproducing the anatomy of the vocal tract during the articulation of sustained oral vowels in all individual variants, depending on the personal features of the patient's voice under both normal and pathological conditions;
— reflecting the losses in the vocal tract and the radiation impedance of the mouth and nose orifices;
— introducing an additional parameter which would express in a quantitative manner the degree of nasalization determined by the extent of cleft palate.

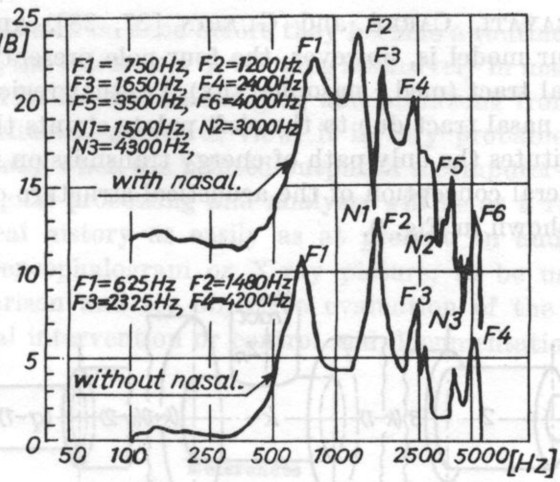

Fig. 8. Example of experimental spectral characteristics of the neutral vowel ɜ measured on an analogue model of the vocal tract without nasalization (lower curve) and with full nasalization (upper curve) [21].
[F_1, F_2,... denote formants, and N_1, N_2,... – antiformants

An accurate modelling of the vocal tract using computer simulation is the subject and goal of research work which is currently being carried on in many countries. Particular areas and aims of the research, however, may be different. The studies by FANT and his associates [5-7, 36], and also the recent works by MRAYATI, CARRÉ and GUERIN in France [27, 28] concentrate mainly on the analysis of relations which exist between the vowel formant frequencies and bandwidths on the one hand, and the acoustical characteristics of the physical model on the other hand, including the boundary effects related to the subglottal system and glottis, losses in the vocal cavity walls and the lip termination. The primary aim of FLANAGAN and co-authors [9-11, 17], on the other hand, has been to develop a working model of the vocal tract and the larynx source for the synthesis of speech in man-machine voice communication systems.

The present state of art in vocal tract modelling and some of the problems involved have recently been discussed in detail by WAKITA and FANT [36]. The conclusion is that much research work must still be done before an improved and standardized model of the vocal tract can be established. In the meantime, however, an approach has been made in the Department of Cybernetic Acoustics, IFTR-PAS, toward a simulative model of the vocal tract including the effect of nasalization due to the cleft palate (cf. KACPROWSKI [19, 20]). The model, which is intended for clinical application in computer-aided acoustical diagnostics in laryngology and phoniatrics, is based on the distributed element transmission line representation of the vocal tract, approximated in the form of a concatenation of n elementary T - type four-poles, as proposed by FLANAGAN [8], and resembles to some extent the recent French model described by MRAYATI, CARRÉ and GUERIN [27, 28]. An essential feature and novelty of our model is, however, the four-pole presentation of the bifurcation of the vocal tract (med.: nasopharynx), i.e. that region where the input impedance of the nasal tract due to the cleft palate shunts the pharynx-mouth tract which constitutes the only path of energy transmission in purely oral articulation. The general conception of the acoustical structure of the latter model is schematically shown in Fig. 9.

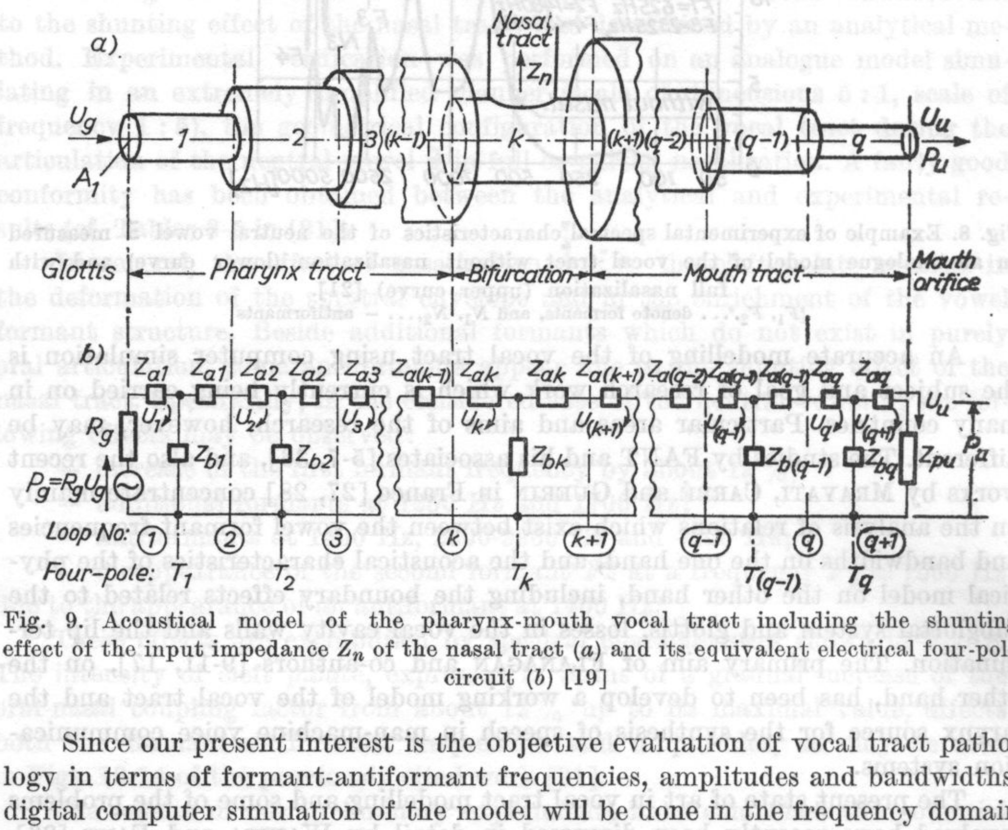

Fig. 9. Acoustical model of the pharynx-mouth vocal tract including the shunting effect of the input impedance Z_n of the nasal tract (a) and its equivalent electrical four-pole circuit (b) [19]

Since our present interest is the objective evaluation of vocal tract pathology in terms of formant-antiformant frequencies, amplitudes and bandwidths, digital computer simulation of the model will be done in the frequency domain only, in order to allow the frequency resolution up to a few Hz for the sake of accuracy, in the case of sustained voiced speech sounds with glottal excitation.

6. Conclusions

The aim of the present paper has been to indicate the potential possibilities of using acoustical diagnostic methods for objective evaluation of larynx and vocal tract pathology, and to present the actual state of research in this domain. Although some of the experimental methods and systems have recently been applied with success in laryngology and phoniatrics, a great deal of research work must still be done before they become a routine in clinical practice. Since, however, the research is developing intensively in many countries and the results obtained so far are encouraging and promising from both scientific and practical, i.e. clinical, points of view, it is very probable, if not certain, that the time is coming when the printed output of a computer-aided diagnostic system of speech signal processing and analysis will be as a rule inserted into the patient's medical history as easily as at present an audiogram, electrocardiogram, electroencephalogram or X-ray picture, to be used later on for year-to-year comparison and for objective evaluation of the effectiveness of the therapy, surgical intervention or control and documentation of the rehabilitation process.

References

[1] J. W. van den Berg, *Myoelastic-aerodynamic theory of voice production*, J. Speech Hearing Res., **1**, 227-244 (1959).

[2] J. W. van den Berg, J. T. Zantema, P. Doornenbal Jr., *On the air resistance and the Bernoulli effect of the human larynx*, J. Acoust. Soc. Am., **29**, 626-631 (1957).

[3] S. B. Davis, *Computer evaluation of laryngeal pathology based on inverse filtering of speech*, Speech Comm. Res. Lab., Monograph No 13, Santa Barbara, California (1976).

[4] G. Fant, *Acoustic theory of speech production*, Mouton and Co., s'-Gravenhage, 1960.

[5] G. Fant, *Vocal tract wall effects, losses and resonance band widths*, STL-QPSR, 2-3, 28-52 (1972).

[6] G. Fant, *Vocal tract area and length perturbations*, STL-QPSR, 4, 1-14 (1975).

[7] G. Fant, L. Nord, P. Branderud, *A note on the vocal tract wall impedance*, STL-QPSR, 4, 13-20 (1976).

[8] J. L. Flanagan, *Speech analysis, synthesis and perception*, Springer Verlag, Berlin – Heidelberg – New York, 1st edition 1965, 2nd edition 1972.

[9] J. L. Flanagan, *Focal points in speech communication research*, Report 21-G-3, Proc. 7th ICA Congress, Budapest, 1971.

[10] J. L. Flanagan, L. L. Landgraf, *Self-oscillating source for vocal-tract synthesizers*, IEEE Trans. Audio Electroacoustics, AU-16, 57-64 (1968).

[11] J. L. Flanagan, K. Ishizaka, K. Shipley, *Synthesis of speech from a dynamic model of the vocal cords and vocal tract*, Bell System Techn. Journal, **54**, 485-506 (1975).

[12] R. Gubrynowicz, W. Mikiel, P. Żarnecki, *Evaluation de l'état pathologique des cordes vocales d'après l'analyse des variations du fondamental*, 8-èmes Journées d'Etudes du Groupe de la Communication Parlée, Vol. 1, 22-27, Institute de Phonétique, Aix-en-Provence, France, 1977.

[13] R. Gubrynowicz, W. Mikiel, P. Żarnecki, *An acoustical method for evaluating the larynx source in the case of vocal folds pathology*, Archives of Acoustics (in print).

[14] R. Gubrynowicz, W. Mikiel, P. Żarnecki, *Acoustical analysis for evaluation of laryngeal disfunction in the case of vocal cord paralysis*, Speech Analysis and Synthesis, Vol. 5, IFTR-PAS, PWN, Warszawa (in press).

[15] K. Ishizaka, *On model of the larynx*, J. Acoust. Soc. Am., **22**, 293-294 (1968).

[16] K. Ishizaka, M. Matsuidara, *What makes the vocal cords vibrate*? Proc. 6th ICA Congress, Report B-1-3, Tokyo 1968.

[17] K. Ishizaka, J. L. Flanagan, *Synthesis of voiced sounds from a two-mass model of the vocal cords*, Bell System Techn. Journal, **51**, 1233-1268 (1972).

[18] J. Kacprowski, *Physical models of the larynx source*, Archives of Acoustics, **2**, 1, 47-70 (1977).

[19] J. Kacprowski, *A simulative model of the vocal tract including the effect of nasalization*, Archives of Acoustics, **2**, 4, 235-255 (1977).

[20] J. Kacprowski, *An acoustical model of the vocal tract for diagnostics of cleft palate*, Speech Analysis and Synthesis, Vol. 5, IFTR-PAS, PWN Warszawa (in print).

[21] J. Kacprowski, W. Mikiel, A. Szewczyk, *Acoustical modelling of cleft palate*, Archives of Acoustics, **1**, 2, 137-158 (1976).

[22] Y. Koike, J. D. Markel, *Application of inverse filtering for detecting laryngeal pathology*, Annals Otol. Rhinol. Laryngol., **84**, 117-124 (1975).

[23] J. D. Markel, A.H. Gray, *Linear prediction of speech*, Springer Verlag, Berlin-Heidelberg-New York 1976.

[24] W. Mikiel, P. Żarnecki, *A computer system for measuring and processing of acoustical data*, Proc. XXIII Open Seminar on Acoustics, Polish Acoust. Soc., 70-72, Wisła 1976 (in Polish, abstract in English).

[25] W. Mikiel, R. Gubrynowicz, W. Hagmajer, *Intonograph — a system for measuring and visualization of the intensity and melody of speech*, Proc. XXIX Open Seminar on Acoustics Polish Acoust. Soc., 120-124, Gdańsk—Władysławowo 1977 (in Polish, abstract in English).

[26] R.L. Miller, *Nature of the vocal cord wave*, J. Acoust. Soc. Am., **31**, 667-677 (1959).

[27] M. Mrayati, R. Carré, *Relations entre la forme du conduit vocal et les caractéristiques acoustiques des voyelles françaises*, E.N.S. d'Électronique et de Radioéléctricité de Grenoble (unpublished report), 1975.

[28] M. Mrayati, B. Guerin, *Études des caractéristiques acoustiques des voyelles françaises par simulation du conduit vocal avec pertes*, Revue d'Acoustique, **9**, 36, 18-32 (1976).

[29] W. H. Perkins, *Vocal function: a behavioural analysis*, Handbook of Speech Pathology and Audiology, Appleton Century Croft, New York 1971.

[30] W. Tłuchowski, W. Mikiel, A. Komorowska, E. Widlicka, *An approach to the correlation of bioelectric and phonospectroscopic investigations in different types of cleft palate operations in children*, Otolaryngologia Polska, **28**, 6a, 70-77 (1974) (in Polish, abstract in English).

[31] W. Tłuchowski, W. Mikiel, A. Niedźwiecki, A. Komorowska, *Phonospectroscopic investigations in one-sided and two-sided paresis of laryngeal nerves*, Diary of the 29th Meeting of Polish Oto-Laryngologists, Białystok 1974, 359-363 (in Polish).

[32] W. Tłuchowski, J. Kacprowski, W. Mikiel, A. Niedźwiecki, A. Komorowska, *Une méthode nouvelle d'analyse phonospectroscopique utilisée dans la phoniatrie clinique*, Proc. XVI Intern. Congress on Logopedics and Phoniatrics, Interlaken 1974, Abstract in: Folia Phoniatrica, **26**, 3, 227 (1974).

[33] W. Tłuchowski, J. Kacprowski, W. Mikiel, A. Niedźwiecki, A. Komorowska, *Clinical phonospectroscopy — a preliminary report*, Otolaryngologia Polska, **29**, 3, 251-259 (1975) (in Polish, abstract in English).

[34] W. Tłuchowski, W. Mikiel, R. Gubrynowicz, W. Komorowska, *Acoustical examination of glottal movements in recurrent laryngeal palsy*, Proc. 4th Congress of the Union of European Phoniatricians, Wrocław 1975.

[35] W. Tłuchowski, W. Mikiel, B. Maniecka-Aleksandrowicz, A. Szewczyk, *Phonospectroscopic investigations of physiological phonation activity of the larynx in pre-school children*, Proc. VI Days of Children Otolaryngology, Łódź 1977, Otolaryngologia Polska, **31**, Supplement, 150-151 (1977) (in Polish, abstract in English).

[36] H. Wakita, G. Fant, *Toward a better model of the vocal tract*, STL-QPSR, 1, 9-29 (1978).

Received on May 4, 1979