ARCHIVES OF ACOUSTICS
1, 2, 137-158 (1976)

ACOUSTICAL MODELLING OF CLEFT PALATE

JANUSZ KACPROWSKI, WŁADYSŁAW MIKIEL,
ALICJA SZEWCZYK*

Institute of Fundamental Technological Research
Polish Academy of Sciences (Warszawa)

The subject of the present paper is the theoretical analysis of the influence of cleft palate upon the transmission characteristics of the vocal tract in glottal excitation. Starting with a simplified anatomical model of the human vocal system and using the graphical analysis method, we determined the pole-zero distribution of the transfer function $K(\omega)$ which describes the formant-antiformant structure of oral vowels nasalized due to the shunting effect of the nasal channel. The experimental investigations, performed with a specially designed and constructed acoustic analogue model of the speech organ in 5 : 1 dimension scale and, consequently, in 1 : 5 frequency scale, confirmed the validity of the assumptions and the results of the theoretical analysis. It was proved that the spectral analysis of oral vowels, in pathological and post-operative states of cleft palate, affords information which specifies the changes of the acoustic structure of the speech organ due to anatomical disorders. This information offers new facilities in objective phoniatric diagnostics as well as in the control of the process of medical rehabilitation in cleft palate cases, by means of acoustical methods based on the spectral analysis of the speech signal.

1. Introduction

The cleft palate (med.: *palatoschisis*) is a frequently-occurring developmental defect of the human respiratory and articulatory system which causes considerable speech disorders. The laryngological treatment consists in complicated and multistage surgical interventions, usually complemented by a long-term rehabilitation procedure. In view of the rather frequent occurrence of this defect, which averages — depending on demographic conditions — one case for several hundreds to one thousand of births, the treatment of cleft palate is a serious medical problem of considerable social as well as scientific importance. At present it constitutes one of the fundamental research trends, in jawbone and plastic surgery and in modern phoniatry.

From the physiological point of view the cleft palate consists of a discontinuity of tissue taking the form of a split in the palatal bone (med.: *os palatinum*) of the hard palate (med.: *palatum*) which separates the mouth cavity

*) At present: Central Clinical Hospital of the Medical Academy in Warszawa, Phoniatric Centre.

from the nasal one. In other cases it consists in the restraint of the motive functions or even in the atrophy of the soft palate (med: *velum*) which acts as a valve controlling the degree of coupling between the mouth cavity and the nasal cavity in the vocal system [1]. In any case, the direct effect of these pathological disorders is an anomaly in the proper cooperation of the mouth channel and the nasal channel due to the permanent and uncontrolled acoustic coupling between them, which in turn causes the forced nasalization of the originally non-nasalized speech sounds and thus distorts the correct articulation.

Since any change in the anatomical structure of the vocal tract due to cleft palate results in a measurable variation of the acoustic transmittance, which determines the pole-zero distribution of the transfer function $K(\omega)$, one may presume that the spectral analysis of the speech sounds in pathological and postoperative cases should yield essential information which could be used in objective evaluation of the actual state of cleft palate or in objective control of the rehabilitation process.

Acoustic methods of spectral analysis have recently been applied with success in clinical practice to the phoniatric diagnosis of cleft palate [8]. The subject of the present work is the theoretical analysis of the influence of the anatomical defect existing in the soft palate (med.: *palatischisis molle*) upon the transmission characteristics of the vocal tract which determine the formant structure of the speech sounds with glottal excitation, particularly oral vowels. The theoretical analysis is followed by the experimental verification of the results by means of model investigations. Finally, some conclusions are formulated which provide a method of diagnosis based upon the phonospectroscopic analysis of the speech signal segments corresponding to oral vowels nasalized due to a cleft palate.

2. Theoretical analysis

Figure 1 represents the simplified acoustic model of the human vocal system in glottal excitation and its impedance-type analogous circuit. The simplification introduced for the anatomical structure of the articulatory effectors consists of representing the pharynx cavity $P(l_p, S_p)$, the mouth cavity M (l_m, S_m) and the nasal cavity $N(l_n, S_n)$, each as a cylindrical tube of length l and constant cross-sectional area S. The system as a whole is excited by the larynx tone generator of internal impedance Z_g, producing a particle velocity v_g. The resulting particle velocity $v(0)$ of the sound wave at the observation point 0, situated at a distance d from the common plane of the mouth and noise outlets, which have radiation impedances equal to Z_{rm} and Z_{rn}, respectively, may be generally expressed in the form of the vector sum (1). Notice that the phase shifts are disregarded in view of the geometric symmetry of the radiating system, hence

$$v(0) = F_{m0}(f)\vec{v}_m + F_{n0}(f)\vec{v}_n, \tag{1}$$

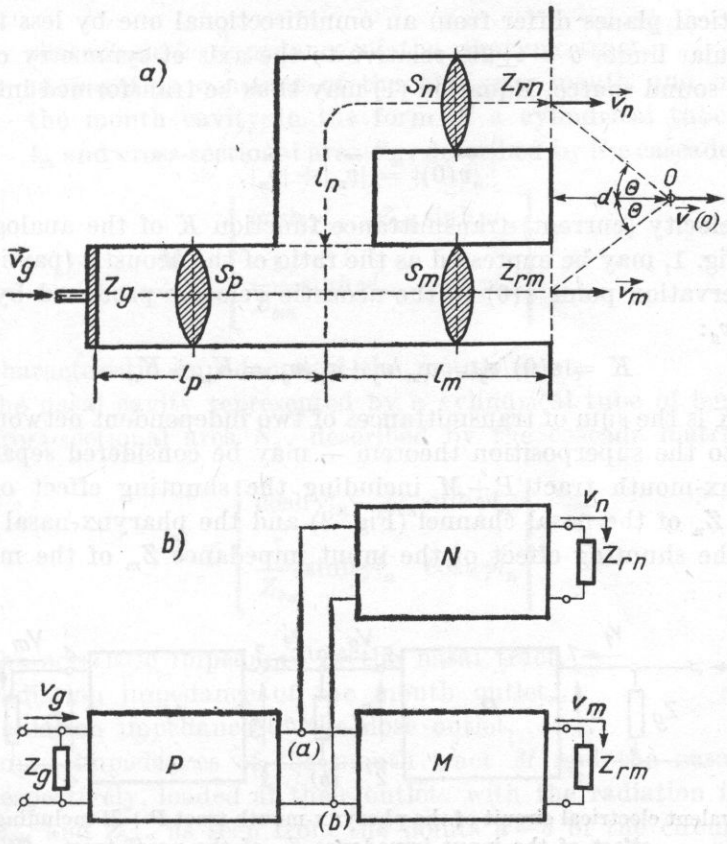

Fig. 1. Simplified acoustic model of the speech organ in glottal excitation (a) and its equivalent electrical circuit (b)

where $\vec{v}_m$ and $\vec{v}_n$ denote acoustic particle velocities of the sound waves radiated by the mouth and nose outlets, respectively, $F_{m\theta}(f)$ and $F_{n\theta}(f)$ – frequency-dependent directional radiation characteristics of the mouth and nose outlets, respectively.

Within the considered frequency range the condition $r_0/\lambda < 0.25$ is fulfilled, where r_0 stands for the radius of the mouth or nose outlet and λ denotes the wavelength. Thus both outlets may be treated as sources of spherical waves. In this connection

$$F_{m\theta}(f) = F_{n\theta}(f) \approx \text{const} = 1, \tag{2}$$

the more so since at a distance $d = 20$ cm, and with symmetrical localization of the observation point 0 with regard to mouth and nose outlets, the radiation angle $\theta = 7°$. Approximation (2) is justified by the results of FLANAGAN's [4] model investigations which showed that, for frequencies up to 5000 Hz, the directional radiation characteristics of the human head in both horizontal

and in vertical planes differ from an omnidirectional one by less than ± 1 dB within angular limits $\theta = \pm 20°$ relative to the axis of symmetry of the outlet acting as a sound source. Equation (1) may thus be transformed into the scalar form:

$$|\vec{v}(0)| = |\vec{v}_m| + |\vec{v}_n|. \tag{3}$$

The velocity (current) transmittance function K of the analogous circuit, shown in Fig. 1, may be expressed as the ratio of the acoustic (particle) velocity at the observation point $v(0)$ to the acoustic velocity produced by the larynx generator v_g:

$$K = v(0)/v_g = v_m/v_g + v_n/v_g = K_m + K_n. \tag{4}$$

Thus K is the sum of transmittances of two independent networks which — according to the superposition theorem — may be considered separately, viz.: the pharynx-mouth tract $P+M$ including the shunting effect of the input impedance Z_n of the nasal channel (Fig. 2) and the pharynx-nasal tract $P+N$ including the shunting effect of the input impedance Z_m of the mouth cavity (Fig. 3).

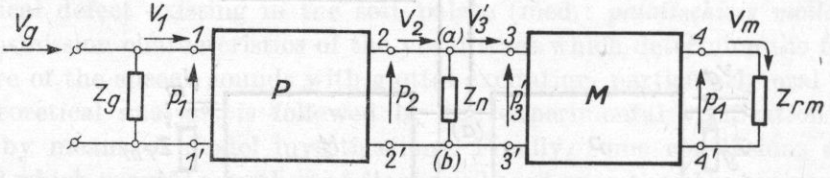

Fig. 2. Equivalent electrical circuit of the pharynx-mouth tract $P+M$ including the shunting effect of the input impedance Z_n of the nasal tract

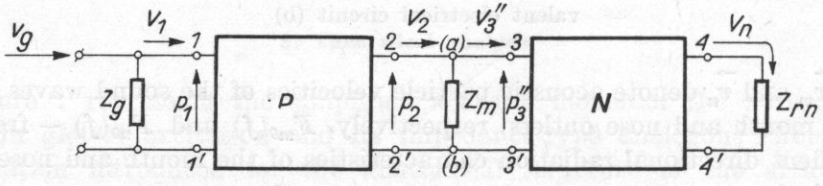

Fig. 3. Equivalent electrical circuit of the pharynx-nasal tract $P+N$ including the shunting effect of the input impedance Z_m of the mouth tract

In Figs. 2 and 3 the following symbols have been used: V_g — the volume velocity produced by the larynx generator, Z_g — the internal impedance of the larynx generator, P — the pharynx cavity in the form of a cylindrical tube of length l_p and cross-sectional area S_p, described by the cascade (chain) matrix $\mathbf{P}$,

$$\mathbf{P} = \begin{bmatrix} \cosh \gamma l_p & Z_{0p} \sinh \gamma l_p \\ \dfrac{1}{Z_{0p}} \sinh \gamma l_p & \cosh \gamma l_p \end{bmatrix},$$

where

Z_{0p} – characteristic impedance of the pharynx tract,

$\gamma = \alpha + j\beta$ – propagation constant of the pharynx, mouth and nasal tract,

M – the mouth cavity in the form of a cylindrical tube of length l_m and cross-sectional area S_m, described by the cascade matrix **M**,

$$\mathbf{M} = \begin{bmatrix} \cosh \gamma l_m & Z_{0m} \sinh \gamma l_m \\ \frac{1}{Z_{0m}} \sinh \gamma l_m & \cosh \gamma l_m \end{bmatrix},$$

Z_{0m} – characteristic impedance of the mouth tract,

N – the nasal cavity represented by a cylindrical tube of length l_n and cross-sectional area S_n, described by the cascade matrix **N**,

$$\mathbf{N} = \begin{bmatrix} \cosh \gamma l_n & Z_{0n} \sinh \gamma l_n \\ \frac{1}{Z_{0n}} \sinh \gamma l_n & \cosh \gamma l_n \end{bmatrix},$$

Z_{0n} – characteristic impedance of the nasal tract,

Z_{rm} – radiation impedance of the mouth outlet,

Z_{rn} – radiation impedance of the nose outlet,

Z_m, Z_n – input impedances of the mouth tract M and the nasal tract N, respectively, loaded at their outlets with the radiation impedances Z_{rm} and Z_{rn}, as seen from the points $a-b$ of the circuit,

$$Z_m = Z_{0m} \frac{Z_{rm} + Z_{0m} \operatorname{tgh} \gamma l_m}{Z_{0m} + Z_{rm} \operatorname{tgh} \gamma l_m},$$

$$Z_n = Z_{0n} \frac{Z_{rn} + Z_{0n} \operatorname{tgh} \gamma l_n}{Z_{0n} + Z_{rn} \operatorname{tgh} \gamma l_n},$$

V_m, V_n – volume velocities in the mouth and nose outlets, respectively,

$V_1, V_2, \ldots; p_1, p_2, \ldots$ – volume velocities and acoustic pressures at different points of the circuit.

The volume velocity (current) transmittance function K_m of the pharynx--mouth tract $P+M$ including the shunting effect of the nasal tract N, expressed as the ratio of the volume velocity in the mouth outlet V_m to the volume velocity produced by the larynx generator V_g may be represented – in agreement with the notation used in Fig. 2 – in the form of the following product:

$$K_m = \frac{V_m}{V_g} + \frac{V_m}{V_3'} \frac{V_3'}{V_2} \frac{V_2}{V_1} \frac{V_1}{V_g}. \tag{5}$$

Using the general rules and methods of linear electric network analysis and some elementary algebraic operations, one can transforme equation (5) to

$$K_m(f) = \frac{1}{\cosh\gamma l_m + \frac{Z_{rm}}{Z_{0m}}\sinh\gamma l_m}\,\frac{Z_n}{Z_n+Z_m}\times$$

$$\times\frac{1}{\cosh\gamma l_p + \frac{Z_2}{Z_{0p}}\sinh\gamma l_p}\,\frac{Z_g}{Z_g+Z_{0p}\dfrac{Z_2+Z_{0p}\operatorname{tgh}\gamma l_p}{Z_{0p}+Z_2\operatorname{tgh}\gamma l_p}} \tag{6}$$

$$= \frac{\cosh\gamma' l_m}{\cosh(\gamma+\gamma')l_m}\,\frac{Z_n}{Z_n+Z_m}\,\frac{\cosh\gamma'' l_p}{\cosh(\gamma+\gamma'')l_p}\,\frac{Z_g}{Z_g+Z_{0p}\dfrac{Z_2+Z_{0p}\operatorname{tgh}\gamma l_p}{Z_{0p}+Z_2\operatorname{tgh}\gamma l_p}},$$

where

$Z_2 = Z_nZ_m/(Z_n+Z_m)$ — equivalent impedance of the input impedances Z_m and Z_n of the mouth tract and the nasal tract, respectively, in parallel connection,

$$\gamma' l_m = \operatorname{tgh}^{-1}(Z_{rm}/Z_{0m}), \qquad \gamma'' l_p = \operatorname{tgh}^{-1}(Z_2/Z_{0p}).$$

The last substitutions express the radiation impedance Z_{rm} of the mouth tract and the load impedance Z_2 of the pharynx tract as formal modifications of the propagation constant γ of each tract.

Using the same procedure in the case of the pharynx-nasal tract $(P+N)$ in Fig. 3, one can express its volume velocity transmittance function $K_n = V_n/V_g$ in the form

$$K_n(f) = \frac{1}{\cosh\gamma l_n + \frac{Z_{rn}}{Z_{0n}}\sinh\gamma l_n}\,\frac{Z_m}{Z_m+Z_n}\times$$

$$\times\frac{1}{\cosh\gamma l_p + \frac{Z_2}{Z_{0p}}\sinh\gamma l_p}\,\frac{Z_g}{Z_g+Z_{0p}\dfrac{Z_2+Z_{0p}\operatorname{tgh}\gamma l_p}{Z_{0p}+Z_2\operatorname{tgh}\gamma l_p}}$$

$$= \frac{\cosh\gamma''' l_n}{\cosh(\gamma+\gamma''')l_n}\,\frac{Z_m}{Z_m+Z_n}\,\frac{\cosh\gamma'' l_p}{\cosh(\gamma+\gamma'')l_p}\,\frac{Z_g}{Z_g+Z_{0p}\dfrac{Z_2+Z_{0p}\operatorname{tgh}\gamma l_p}{Z_{0p}+Z_2\operatorname{tgh}\gamma l_p}}, \tag{7}$$

where, as before,

$$Z_2 = \frac{Z_mZ_n}{Z_m+Z_n}, \qquad \gamma'' l_p = \operatorname{tgh}^{-1}(Z_2/Z_{0p})$$

and, consequently,

$\gamma''' l_p = \operatorname{tgh}^{-1}(Z_{rn}/Z_{0n})$ — corrected propagation constant of the nasal tract including its radiation impedance Z_{rn}.

The transmittance functions of both channels, i.e., $K_m(f)$ of the pharynx--mouth tract and $K_n(f)$ of the pharynx-nasal tract, are complicated when account is taken of mutual shunting effects. The determination of the pole-zero distributions by means of equations (6) and (7) is thus a rather difficult mathematical task, particularly because most acoustic parameters that occur in these equations are compound. For example, the radiation impedances Z_r of the mouth and nose outlets, even when approximated by the most simple model in the form of a circular piston in an infinite plane baffle, are described be the expression of the type [7],

$$Z_r = \frac{\varrho c}{S}\left[\left(1-\frac{J_1(2\beta r)}{\beta r}\right)+j\left(\frac{K_1(2\beta r)}{2(\beta r)^2}\right)\right], \tag{8}$$

where $\beta = \omega/c = 2\pi/\lambda$, r is the piston radius, $S = \pi r^2$ — the piston area, $J_1(x)$ — first order Bessel function of the first kind, $K_1(x)$ — first order Struve function and ϱc — specific acoustic impedance of the medium. For frequencies such that the condition $\beta r \ll 1$ is fulfilled, the radiation impedance (8) may be approximated as

$$Z_r \approx \frac{\varrho c}{S}\left[\frac{(\beta r)^2}{2}+j\frac{8}{3\pi}(\beta r)\right]. \tag{8a}$$

An equally compound form is displayed by the internal impedance $Z_g = R_g + j\omega M$, whose dominant real or resistive component $R_g \gg \omega M_g$ is a function of the geometrical dimensions of the orifice of the vocal cords, i.e. its length l, width w and depth d, as well as of the volume velocity flow V_g through the orifice [2],

$$R_g(t) \approx \frac{12\mu d}{lw^3(t)}+0.875\frac{\varrho V_g(t)}{2[lw(t)]^2}, \tag{9}$$

where μ is the viscosity coefficient of the medium, i.e. the air flowing through the orifice.

Furthermore, the lossy or dissipative parameters of the vocal tract, i.e. the acoustic viscous resistance per unit length $R_a(f)$ and the acoustic conductance per unit length $G_a(f)$ due to heat conduction losses, are, in general, frequency dependent. Consequently, the real component α of the propagation constant $\gamma = \alpha + j\beta$ of the vocal tract cannot be considered to be a constant and frequency independent quantity, since

$$\alpha = \alpha_R + \alpha_G \approx \frac{R_a}{2}\sqrt{C_a/L_a}+\frac{G_a}{2}\sqrt{L_a/C_a} = f(\omega). \tag{10}$$

For the above-mentioned reasons the determination of the pole-zero distribution of the transmittance functions $K_m(f)$, formula (6), and $K_n(f)$, formula (7), in the complex frequency plane $s = \sigma + j\omega$ can be achieved only by numeri-

cal methods using computer technique. Such a procedure, although possible and sometimes used, proved to be neither necessary nor purposeful, if – as in this case – the task consists of calculating only the characteristic frequencies of the vocal tract which describe the spectral structure of vowel sounds including the effect of nasalization due to the cleft palate. For this purpose it is more convenient to introduce further, rationally motivated, simplifications of the acoustic structure of the anatomical model of the human vocal system in Fig. 1, and to transform its equivalent electrical circuit into a shape suitable for functional analysis. These simplifications are based on the following assumptions:

(a) The acoustic tubes which simulate the pharynx (P), mouth (M) and nasal (N) tracts and, consequently, the corresponding electrical four-poles, are considered as nondissipative (lossless) systems: $R_a \approx 0$, $G_a \approx 0$. In this case the attenuation constant α in (10) equals zero, the propagation constant $\gamma = j\beta$ is an imaginary quantity, whereas the characteristic impedance Z_0 of the considered acoustic tube having the cross-sectional area S is a real quantity: $Z_0 = \varrho c/S$.

(b) The radiation impedances of the mouth and nose outlets are considered as being equal to zero: $Z_{rm} = Z_{rn} = 0$. Consequently, the output terminals of the four-poles M and N in the corresponding equivalent electrical circuits are short-circuited.

(c) The larynx source is considered as a constant current generator, with an internal impedance $Z_g = \infty$, which delivers to the vocal tract a constant volume velocity, i.e. $V_g = \text{const}$ and is indepedent of the load impedance, in this case the input impedance Z_{in} of the vocal tract; this assumption is fully motivated since $|Z_g| \gg |Z_{in}|$.

The simplifying assumption (a) has negligible influence upon the accuracy of the pole frequency values of the transmittance functions, particularly those of the pharynx-mouth tract function whose attenuation constant $\alpha = \alpha_R + \alpha_G$, formula (10), is rather small, being of the order $5 \cdot 10^{-4}$ at $f = 500$ Hz, and increases with the square root of frequency [3, 5]. The influence of the radiation impedance Z_{rm} on the formant frequencies of the pharynx-mouth tract (simplifying assumption (b)) has been investigated using a theoretical model in the form of a cylindrical tube of length $l = 17$ cm and cross-sectional area $S = 5$ cm^2. It can be shown (cf. e.g. [5], p. 61–63) that loading the outlet of the tube with the radiation impedance Z_r, formula (8a), increases its effective length by $\Delta l' = 8r/3\pi$, where r is the orifice radius. It thus lowers the first and second formant frequencies from the values $F_1 = 500$ Hz and $F_2 = 1500$ Hz, calculated for $Z = 0$, to $F_1 = 470$ Hz and $F_2 = 1410$ Hz, respectively, it is by only about 6%. Using the same theoretical model of the vocal tract one can show that the condition $|Z_g| \gg |Z_{in}|$, supporting the simplifying assumption (c), is fulfilled for a broad frequency range, except in the vicinity of the first formant F_1, where $|Z_{in}|$ reaches its maximum value $|Z_{in}|_{\max} \approx 80 \cdot 10^5$ MKS acoustic ohms, and then decreases with further increase of frequency to the asymptotic

value $|Z_{in}| = |Z_r| \approx 8 \cdot 10^5$ MKS acoustic ohms, which is equal to the output radiation impedance. Substitution, for the internal impedance $Z_g = \infty$ of the larynx generator, of the finite value $Z_g = R_g + j\omega M_g = (91 + j\omega\, 6.8 \cdot 10^{-3}) \cdot 10^5$ MKS acoustic ohms, which corresponds to the average geometrical dimensions of the vocal cords orifice of an adult male ($l = 18$ mm, $w = 0.28$ mm, $d = 3$ mm) at mean subglottal pressure $P_s = 8$ cm H_2O, increases the first and second formant frequencies from the theoretical values $F_1 = 500$ Hz and $F_2 = 1500$ Hz, calculated for $Z_g = \infty$, by about $1.4\,\%$ and $1.0\,\%$, respectively, due to the apparent shortening of the effective length $l = 17$ cm of the tube by the value $\Delta l'' = M_g Z_0 c |Z_g|^{-2}$, where

$Z_0 = \varrho c / S$ — characteristic impedance of the tube,

$Z_g = \sqrt{(R_g^2 + \omega^2 M_g^2)}$ — modulus of the larynx generator's impedance,

c — sound velocity in air.

Having introduced the simplifying assumptions (a), (b) and (c), we can represent the equivalent circuits of the pharynx-mouth tract $P + M$ (Fig. 2) and of the pharynx-nasal tract $P + N$ (Fig. 3) in the forms shown in Figs. 4a and 4b.

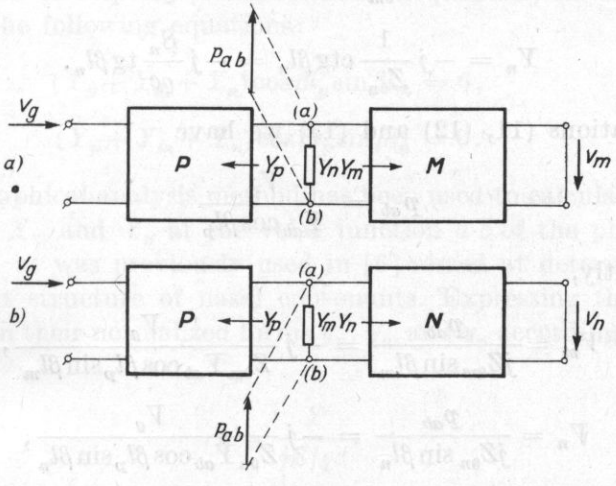

Fig. 4. Simplified equivalent electrical circuits of (a) the pharynx-mouth tract $P + M$ and (b) of the pharynx-nasal tract $P + N$

The velocity or current transmittance functions $K_m = V_m / V_g$ and $K_n = V_n / V_g$ of both tracts can now be determined by elementary methods based on well-known rules of the general theory of linear electric networks adapted to acoustic systems. According to NORTON's theorem, the acoustic pressure (or electric voltage) p_{ab} at the points a-b of the circuti may be expressed as

$$P_{ab} = \frac{V_z}{Y_{ab}}, \tag{11}$$

where

$$V_z = \frac{V_g}{\cos \beta l_p} \tag{12}$$

is the output volume velocity (or output current) of the four-pole P, provided the points a-b are short-circuited, and

$$Y_{ab} = Y_p + Y_m + Y_n \tag{13}$$

is the acoustic admittance of the system, measured between the points a-b, and is equal to the sum of the input admittances of the four-poles P, M and N as seen from the terminals a-b. In view of the previously introduced simplifications we get:

$$Y_p = j \frac{1}{Z_{0p}} \operatorname{tg} \beta l_p = j \frac{S_p}{\varrho c} \operatorname{tg} \gamma l_p, \tag{14a}$$

$$Y_m = -j \frac{1}{Z_{0m}} \operatorname{ctg} \beta l_m = -j \frac{S_m}{\varrho c} \operatorname{tg} \beta l_m, \tag{14b}$$

$$Y_n = -j \frac{1}{Z_{0n}} \operatorname{ctg} \beta l_n = -j \frac{S_n}{\varrho c} \operatorname{tg} \beta l_n. \tag{14c}$$

From equations (11, (12) and (13) we have

$$p_{ab} = \frac{V_g}{Y_{ab} \cos \beta l_p} \tag{15}$$

and, consequently,

$$V_m = \frac{p_{ab}}{j Z_{0m} \sin \beta l_m} = -j \frac{V_g}{Z_{0m} Y_{ab} \cos \beta l_p \sin \beta l_m}, \tag{16}$$

$$V_n = \frac{p_{ab}}{j Z_{0n} \sin \beta l_n} = -j \frac{V_g}{Z_{0n} Y_{ab} \cos \beta l_p \sin \beta l_n}, \tag{17}$$

whence, finally,

$$K_m = \frac{V_m}{V_g} = -j Y_{0m} [Y_{ab} \cos \beta l_p \sin \beta l_m]^{-1}, \tag{18}$$

$$K_n = \frac{V_n}{V_g} = -j Y_{0n} [Y_{ab} \cos \beta l_p \sin \beta l_n]^{-1}, \tag{19}$$

where $Y_{0m} = Z_{0m}^{-1}$ and $Y_{0n} = Z_{0n}^{-1}$.

Identical results could be obtained from the general expressions for K_m in formula (6) and K_n in formula (7) after introducing the approximations

$$\gamma = \alpha + j\beta \approx j\beta, \quad Z_{rm} = Z_{rn} \approx 0, \quad Z_g = \infty,$$

which result from the simplifying assumptions (a), (b) and (c). In point of fact, equation (6) for K_m may be rewritten as follows:

$$K_m \approx \frac{1}{\cos\beta l_m}\,\frac{Y_m}{Y_m+Y_n}\,\frac{1}{\cos\beta l_p + j\,\dfrac{1}{Z_{0p}}\,\dfrac{1}{(Y_n+Y_m)}\sin\beta l_p}$$

$$= \frac{1}{\cos\beta l_m}\,\frac{Y_m}{Y_m+Y_n}\,\frac{1/\cos\beta l_p}{1+\dfrac{Y_p}{Y_n+Y_m}} = -j\,\frac{Y_{0m}\operatorname{ctg}\beta l_m}{\cos\beta l_m\cos\beta l_p\,Y_{ab}} \tag{18a}$$

$$= -jY_{0m}[Y_{ab}\cos\beta l_p\sin\beta l_m]^{-1}.$$

Similarly, from equation (7), after identical transformations, we get

$$K_n \approx -jY_{0n}[Y_{ab}\cos\beta l_p\sin\beta l_n]^{-1}. \tag{19a}$$

The poles of the transmittance functions K_m, formula (18), and K_n, formula (19), of the pharynx-mouth tract and of the pharynx-nasal tract, respectively, which correspond to the formant frequencies, can be obtained by the solution of the following equations:

$$(Y_p+Y_m+Y_n)\cos\beta l_p\sin\beta l_m = 0, \tag{20}$$

$$(Y_p+Y_m+Y_n)\cos\beta l_p\sin\beta l_n = 0. \tag{21}$$

The same graphical analysis method has been used to calculate the network admittances Y_p, Y_m and Y_n at the velar junction a-b of the pharynx, mouth and nasal tracts, as was previously used in [6] aimed at determining the formant-antiformant structure of nasal consonants. Expressing the admittances Y_p, Y_m and Y_n in their normalized forms y_p, y_m and y_n according to the general rule

$$y = \frac{Y}{jS/\varrho c}, \tag{22}$$

and assuming equality of cross-sectional areas of all tracts ($S_p = S_m = S_n$), one can rewrite equations (20) and (21) in purely trigonometrical forms,

$$-\cos\beta(l_p+l_m) = \operatorname{ctg}\beta l_n\cos\beta l_p\sin\beta l_m, \tag{23}$$

$$-\cos\beta(l_p+l_n) = \operatorname{ctg}\beta l_m\cos\beta l_p\sin\beta l_n, \tag{24}$$

which are convenient for graphical analysis. The solutions are shown in Figs. 5 and 6 for the typical case where $l_p = 8.5$ cm, $l_m = 8.5$ cm and $l_n = 12.5$ cm, corresponding to the average geometrical dimensions of the vocal tract of an adult male. The sound velocity in air is taken to be $c = 350$ m/s (moist air at human body temperature, $t = 37°$C).

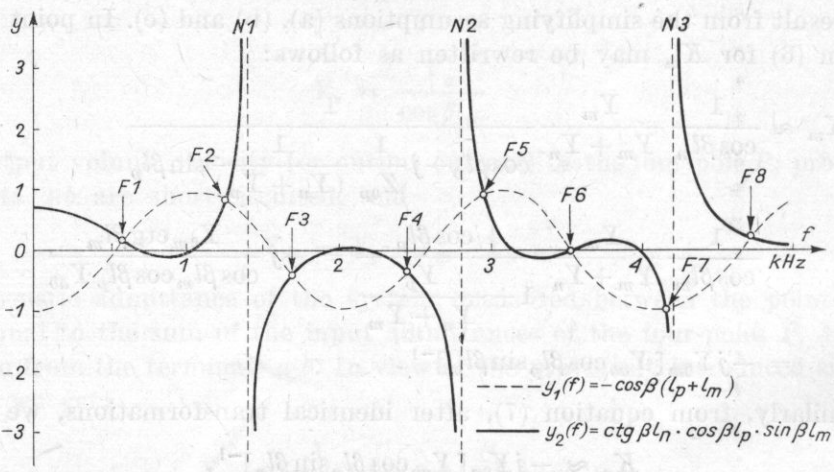

Fig. 5. Determination of the poles FX and zeros NX ($X = 1, 2, 3, \ldots$) of the transmittance function K_m (18) of the pharynx-mouth tract by graphical analysis of (23)

The intersection points of the characteristics

$$y_1(f) = -\cos\beta(l_p + l_m) \quad \text{and} \quad y_2(f) = \operatorname{ctg}\beta l_n \cos\beta l_p \sin\beta l_m$$

in Fig. 5, as well as the intersection points of the characteristics

$$y_3(f) = -\cos\beta(l_p + l_n) \quad \text{and} \quad y_4(f) = \operatorname{ctg}\beta l_m \cos\beta l_p \sin\beta l_n$$

in Fig. 6 correspond to the pole frequencies of the transmittance functions K_m (18) of the pharynx-mouth tract and K_n (19) of the pharynx-nasal tract, respectively, whereas the discontinuity points of the functions $y_2(f) = \pm\infty$

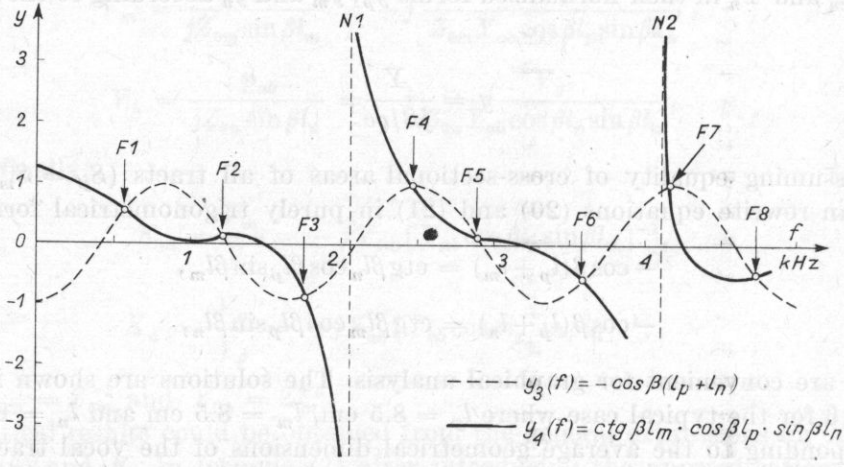

Fig. 6. Determination of the poles FX and zeros NX ($X = 1, 2, 3, \ldots$) of the transmittance function K_n (19) of the pharynx-nasal tract by graphical analysis of (24)

and $y_4(f) = \pm\infty$ determine the zeros of the respective transmittances. The curve $y_1(f) = -\cos\beta(l_p+l_m)$ describes the trivial case of the neutral vowel Э pronounced without any nasalization (no cleft palate: $Y_n = 0$). The intersection points of this curve with the abscissa, that is the points where $y_1(f) = 0$, define the classical formant frequencies

$$F_{(2n+1)} = (2n+1)\cdot\frac{c}{4l} \approx (2n+1)\cdot 515 \text{ Hz}$$

from the general condition

$$\beta l = (2n+1)\cdot\frac{\pi}{2}$$

for $n = 0, 1, 2, \ldots$, where $l = l_p+l_m = 17$ cm. The curve $y_1(f)$ may thus be considered as the reference characteristic corresponding to the articulation of the neutral oral vowel Э in physiologically normal conditions (no cleft palate).

Tables 1 and 2 contain the frequency values of the poles F_x and zeros N_x ($x = 1, 2, 3, \ldots$) of the transmittance functions K_m and K_n describing the pharynx-mouth tract (Fig. 4a) and the pharynx-nasal tract (Fig. 4b), respectively. The pole-zero frequencies were determined by graphical solution of equations (23) and (24) in Figs. 5 and 6.

Table 1. Frequency values of poles F_x and zeros N_x in the transmittance function K_m of the pharynx-mouth tract in Fig. 4a, determined by the graphical method from (23) for $l_p = l_m = 8.5$ cm, $l_n = 12.5$ cm, $c = 350$ m/s

F_x, N_x	F_1	F_2	N_1	F_3	F_4	N_2	F_5	F_6	F_7	N_3	F_8
Hz	575	1250	1400	1700	2450	2800	2960	3520	4150	4200	4720

From the viewpoint of phoniatric diagnostics the most interesting case in the pathological and postoperative states of cleft palate is the general case, when both tracts: the pharynx-mouth and the pharynx-nasal ones, are active simultaneously and shunt one another. In agreement with the previously applied superposition theorem, this case may be considered analytically as the result of summation of the transmittance characteristics (18) and (19), each of which describes a separate channel including the shunting effect of another one.

Table 2. Frequency values of poles F_x and zeros N_x in the transmittance function K_n of the pharynx-nasal tract in Fig. 4b, determined by graphical method from (24) for $l_p = l_m = 8.5$ cm, $l_n = 12.5$ cm, $c = 350$ m/s

F_x, N_x	F_1	F_2	F_3	N_1	F_4	F_5	F_6	N_2	F_7	F_8
Hz	575	1220	1760	2050	2450	2900	3580	4100	4150	4740

The analytical procedure may be interpreted in physical terms as the sum of the acoustic pressures produced at the observation point by the sound waves radiated by the mouth and nose outlets. It is assumed that the point at which the resulting sound pressure is measured lies on the axis of symmetry of the radiating system, and is far enough from both outlets.

It follows that the spectral characteristic of steady-state speech sounds, measured in real conditions, should contain — in principle — all formants and antiformants caused by the nasalization effect and which correspond to the pole-zero distributions of the transmittance functions K_m and K_n of both tracts, as listed in Tables 1 and 2. The theoretical pole-zero distribution of the resulting transmittance function $|K| = |K_m| + |K_n|$ of the complete vocal system, involving both the mouth and the nasal tract, is presented in Table 3.

Table 3. Frequency values of poles F_x and zeros N_x in the transmittance function $K = K_m + K_n$ of the vocal mouth-nasal tract, determined by graphical method in the theoretical case of articulation of the neutral vowel Э, including the nasalization effect due to cleft palate (*palatoschisis molle*), for $l_p = l_m = 8.5$ cm, $l_n = 12.5$ cm, $c = 350$ m/s

F_x, N_x	F_1	F_2	N_1	F_3	N_2	F_4	N_3	F_5	F_6	N_4	F_7
Hz	575	1235	1400	1730	2050	2450	2800	2930	3550	4150	4730

Fig. 7 shows the hypothetical transmittance characteristic $K = f(f)$, which represents the formant-antiformant structure of the neutral vowel Э being nasalized due to the cleft palate effect.

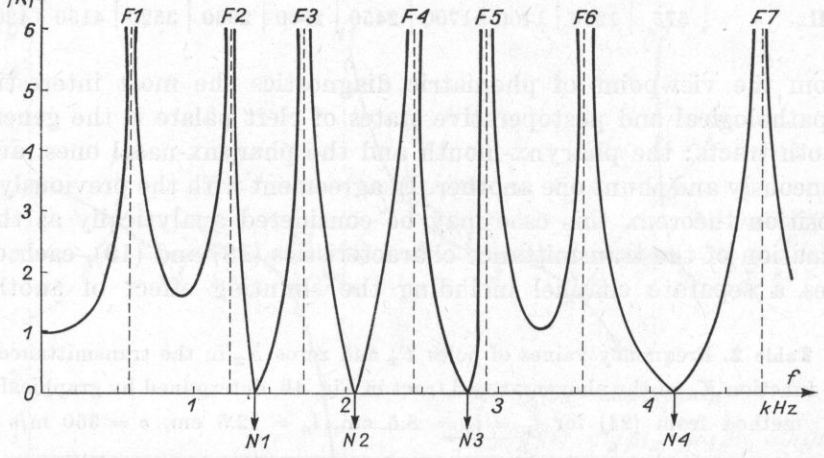

Fig. 7. Distribution of the poles FX an zeros NX ($X = 1, 2, 3, \ldots$) of the hypothetical transmittance function $|K| = |K_m| + |K_n|$ of the vocal mouth-nasal tract, including the mutual effects of both tracts: the mouth channel and the nasal channel

It seems very likely, however, that the speech spectrum envelope measured in real conditions may differ considerably from the theoretical one, due to the mutual cancellation of poles and zeros of both transmittance functions $K_m(f)$ and $K_n(f)$, which are located near one another in the frequency scale.

3. Methodology of experimental investigations

The results of the theoretical analysis have been verified by an experimental investigation performed in a physical system simulating the geometric structure of the anatomical model of the vocal tract (Fig. 1a) in the dimensional scale of 5:1, that is in the frequency scale of 1:5. The choice of the scale factor $k = 5$ was inspired by both metrological and constructional reasons: the lowered frequency range proved to be well adjusted for the electroacoustic parameters of the available measuring equipment, whereas the enlargement of model's geometry improved the accuracy of the system's design and performance.

A simplified schematic diagram of the physical model is shown in Fig. 8. The larynx source *LS* is simulated by a dynamic moving coil loudspeaker, type GDN 16/10, equipped with an acoustic feedback system controlling the

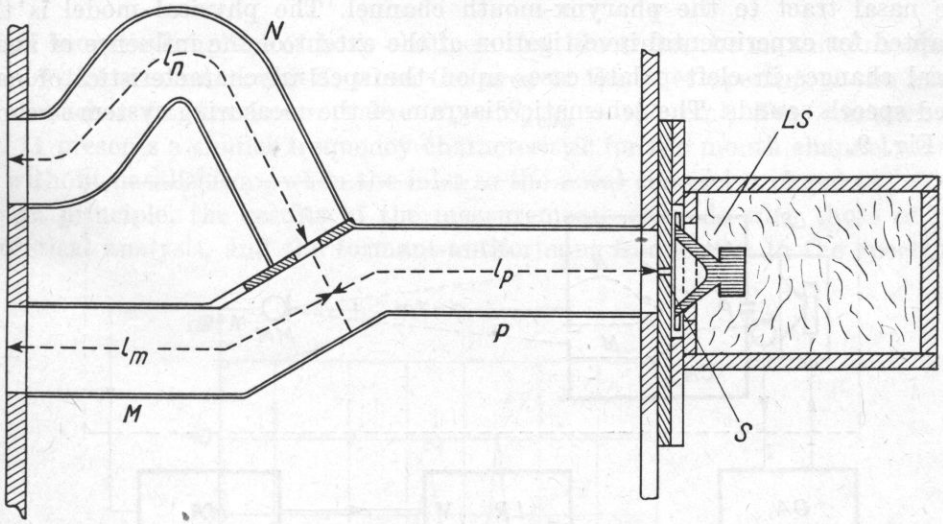

Fig. 8. Simplified schematic diagram of the experimental acoustic model of the speech organ in larynx excitation

LS – larynx source, *S* – slit simulating the vocal cord orifice, *P* – pharynx tract, *M* – mouth tract, *N* – nasal tract

sound pressure level in front of the membrane (p = const). The loudspeaker is installed in a closed and acoustically damped enclosure, whose volume approximates the average volume of the human trachea and bronchi. The larynx source is coupled with the inlet of the pharynx channel *P* through the narrow rectangular slit *S*, whose dimensions ($90 \times 5 \times 15$ mm) correspond to the average

dimensions of the human glottis (length $l = 18$ mm, width $w = 1$ mm, depth $d = 3$ mm) when taking into account the scale factor 5 : 1. The model of the vocal tract itself has the form of a bifurcated system of acoustic tubes of constant cross-sectional area $S = 125\ \text{cm}^2$ and lengths $l_p = 42.5$ cm, $l_m = 42.5$ cm, $l_n = 62.5$ cm, which simulate the pharynx (P), mouth (M) and nasal (N) tract, respectively. The areas of the outlet openings of the mouth and nasal tract are equal to $S_m = 125\ \text{cm}^2$ and $S_n = 50\ \text{cm}^2$, respectively, corresponding to 5 cm² and 2 cm² of the natural scale. Since the model is made of hard and smooth materials (textolite, plexiglas), the acoustic tubes may be considered to be lossless, at least in the frequency range of interest, i.e. up to 1000 Hz.

The human head is simulated by the wooden cubic box whose edge equals 82 cm. The outlet openings of the mouth and nostrils are located vertically in the front wall of the model and the distance between their centres is equal to 22.5 cm, i.e. 4.5 cm in the natural scale. The measuring point 0 lies on the axis of symmetry of the model, at a distance $d = 100$ cm from the plane of the outlet openings. Thus, in real conditions, the measuring microphone should be placed at a distance $d' = 20$ cm in front of the patient's head.

Facilities have been provided for varying the area of the orifice coupling the nasal tract to the pharynx-mouth channel. The physical model is thus adapted for experimental investigation of the extent of the influence of anatomical changes in cleft palate cases upon the spectral characteristics of nasalized speech sounds. The schematic diagram of the measuring system is shown in Fig. 9.

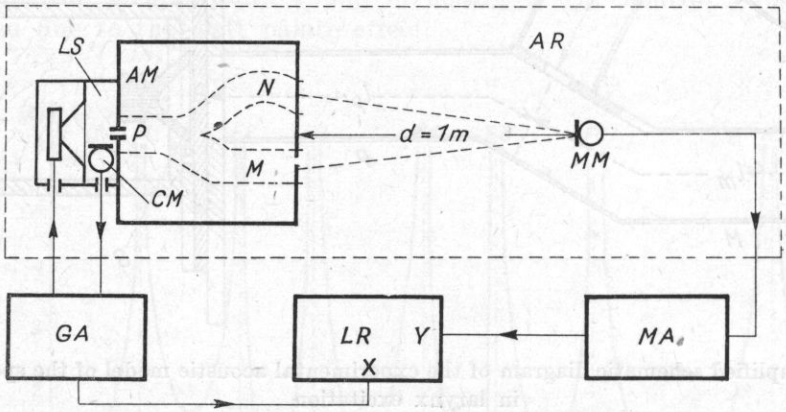

Fig. 9. Schematic diagram of the measuring circuit

AR – anechoic room, *AM* – acoustic model of the speech organ, *P* – pharynx tract, *M* – mouth tract, *N* – nasal tract, *LS* – larynx source, *CM* – control microphone, *MM* – measuring microphone, *MA* – microphone amplifier, *GA* – acoustic generator, *LR* – logarithmic level recorder

4. Results of experiments

Fig. 10 presents the frequency characteristic of the acoustic pressure produced at the measuring point by sound waves radiated from the outlets of both channels: the mouth tract and the nasal tract. The curve corresponds

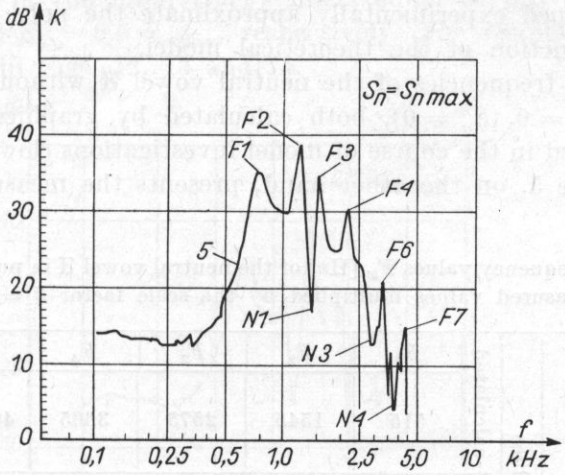

Fig. 10. Radiation frequency characteristic of both model channels: the mouth and nasal tract with full nasalization, corresponding to the maximal inlet area of the nasal channel of $S_n = S_{n\max}$. The frequency scale has been multiplied by 5

to full acoustic coupling of both channels and illustrates the maximal nasalization of the neutral vowel Ы, when the area of the inlet opening of the nasal tract reaches its maximum value $S_n = S_{n\max} = 125$ cm². For comparison, Fig. 11 presents a similar frequency characteristic for the mouth channel alone, i.e. without nasalization, when the inlet to the nasal channel is closed ($S_n = 0$).

In principle, the results of the measurements coincide with those of the theoretical analysis, and the formant-antiformant frequencies in the spectrum

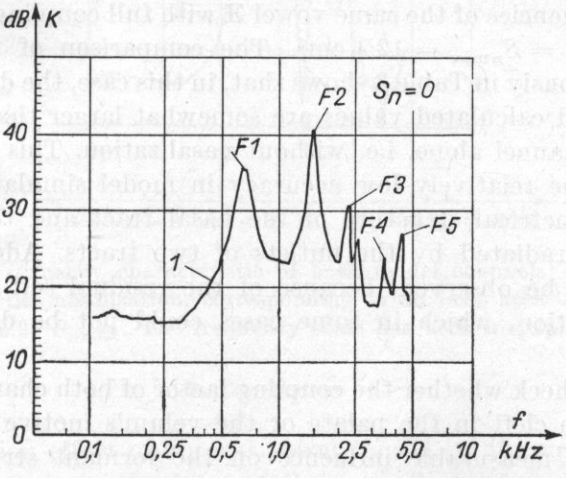

Fig. 11. Radiation frequency characteristic of the model mouth channel without nasalization ($S_n = 0$). The frequency scale has been multiplied by 5

envelope determined experimentally approximate the poles and zeros of the transmittance function of the theoretical model.

The formant frequencies of the neutral vowel ə without nasalization (no cleft palate: $Y_n = 0$, $S_n = 0$), both calculated by graphical method (upper row) and measured in the course of model investigations (lower row), are listed in Table 4. Table 5, on the other hand, presents the measured formant and

Table 4. Formant frequency values F_x [Hz] of the neutral vowel ə in purely oral articulation (measured values multiplied by the scale factor $k = 5$)

Formant frequencies F_x	Values	F_1	F_2	F_3	F_4	F_5
	calculated	515	1545	2575	3605	4635
	measured	625	1500	2400	3500	4600

Table 5. Measured frequency values of formants F_x and antiformants N_x at full acoustic coupling of both model channels: the mouth and the nasal tract, $S_n = S_{n\max}$ (measured values multiplied by the scale factor $k = 5$)

F_x, N_x	F_1	F_2	N_1	F_3	N_2	F_4	N_3	F_5	F_6	N_4	F_7
Hz	700	1250	1400	1700	–	2300	2900	–	3500	4200	4500

antiformant frequencies of the same vowel ə with full coupling of both channels, that is when $S_n = S_{n\max} = 125$ cm². The comparison of these values with those listed previously in Table 3 shows that, in this case, the differences between the measured and calculated values are somewhat larger than those occurring for the mouth channel alone, i.e. without nasalization. This fact may be attributed both to the relatively low accuracy in model simulation of the rather complicated geometrical structure of the nasal tract and to the interference of sound waves radiated by the outlets of two tracts. Additional distorting effects may also be observed, because of the residual self resonances of the model's construction, which in some cases could not be damped effectively enough.

In order to check whether the coupling factor of both channels, representing the extent of the cleft in the palate or the velum's motive capability, exerts an essential and measurable influence on the formant structure of an oral vowel, additional measurements of spectral characteristics of the neutral vowel ə have been performed at a few different values of the coupling factor, corre-

sponding to areas of the nasal tract's inlets equal to $S_n = 0.125\ S_{n\max}$, $S_n = 0.25\ S_{n\max}$, and $S_n = 0.5\ S_{n\max}$, respectively. The results of the measurements are shown in Figs. 12, 13 and 14.

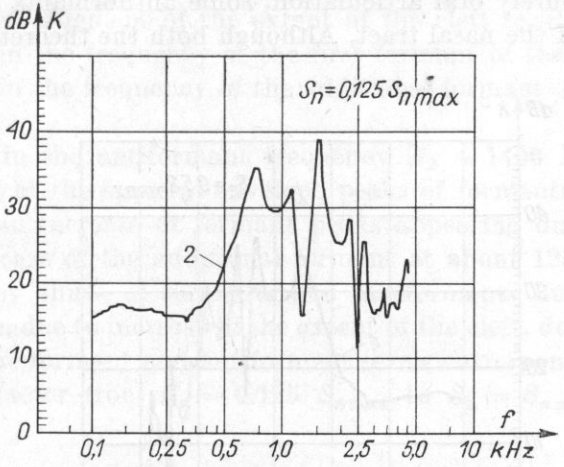

Fig. 12. Radiation frequency characteristic of both model channels: the mouth and the nasal tract with partial nasalization, corresponding to an inlet area of the nasal channel of $S_n = 0.125\ S_{n\max}$. The frequency scale has been multiplied by 5

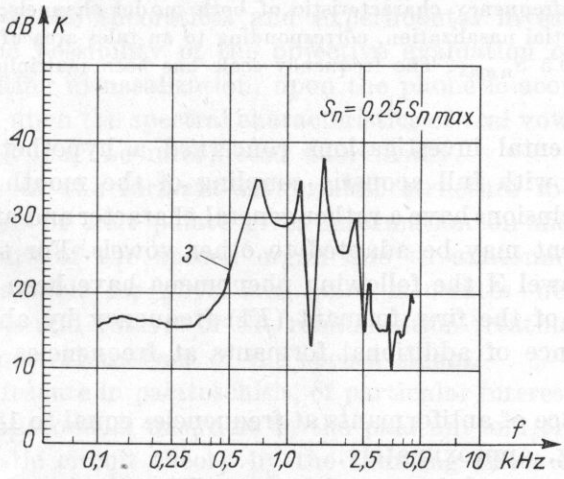

Fig. 13. Radiation frequency characteristic of both model channels: the mouth and the nasal tract with partial nasalization, corresponding to an inlet area of the nasal channel of $S_n = 0.25\ S_{n\max}$. The frequency scale has been multiplied by 5

5. Conclusions

The nasalization effect of an oral vowel caused by the velum's cleft (*palatoschisis molle*) appears in the form of theoretically predicted and experimentally verified modifications of its spectral characteristics. These modifications

consist, generally, in the deformation of the spectral envelope and in the enrichment of the vowel's formant structure. Besides additional formants, which do not exist in purely oral articulation, some antiformants appear due to the shunting effect of the nasal tract. Although both the theoretical considerations

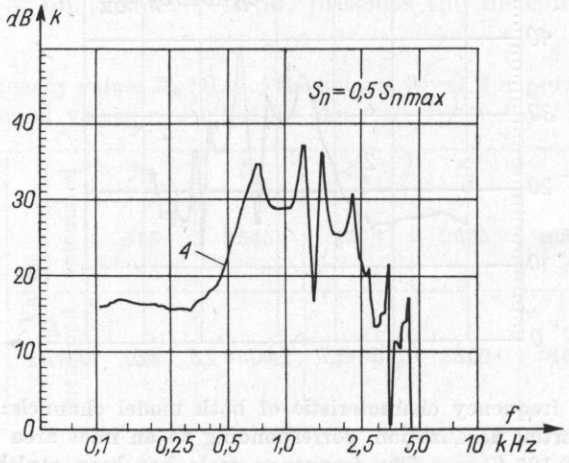

Fig. 14. Radiation frequency characteristic of both model channels: the mouth and the nasal tract with partial nasalization, corresponding to an inlet area of the nasal channel of $S_n = 0.5\ S_{n\max}$. The frequency scale has been multiplied by 5

and the experimental investigations concerned a hypothetical model of the neutral vowel Ə with full acoustic coupling of the mouth and nasal tracts, the resulting conclusions have a rather general character and after similar methodological treatment may be adapted to other vowels. For the particular case of the neutral vowel Ə the following phenomena have been observed:

(a) Increase of the first formant ($F1$) frequency by about 12%.

(b) Appearance of additional formants at frequencies equal to 1250 Hz and 1700 Hz.

(c) Appearance of antiformants at frequencies equal to 1400 Hz, 2800-2900 Hz and 4200 Hz, approximately.

(d) Disappearance of the second formant $F2$ at the frequency $F_2 = 1500$ Hz due to the appearance of a zero in the vocal tract's transmittance function at a frequency $N_1 = 1400$ Hz.

At the same time a slight lowering of the peaks of the spectrum envelope and a corresponding broadening of the bandwidths in the formants $F1$ and $F2$ of the nasalized vowel become quite visible. This phenomenon, which was also observed in the case of nasal consonants (cf. e.g. [6]), has not been interpreted analytically in view of the fact that losses in the vocal tract were *a priori* neglected.

Comparison of the spectral characteristics of the neutral vowel ə at various degrees of nasalization, realized in model investigations by a gradual increase in the inlet area of the nasal tract (cf. Figs. 12, 13 and 14), permit preliminary conclusions that enlargement of the extent of the cleft results in:

(a) Increase in the frequency of the first formant of the neutral vowel ə.

(b) Increase in the frequency of the additional formant occurring at about 1250 Hz.

(c) Increase in the antiformant frequency $N_1 \approx 1400$ Hz.

(d) Lowering of the spectral envelope peaks of formants originating from the oral vowel and increase of formant peaks appearing due to nasalization, especially in the case of the additional formant at about 1250 Hz.

The frequency shifts of formants and antiformants due to the increase of nasalization, i.e. due to increase in the extent of the cleft, do not exceed 10%. The changes of the formant and antiformant levels corresponding to variations of the coupling factor from $S_n = 0.125\ S_{n\max}$ to $S_n = S_{n\max}$ lie within the limits of 5 dB.

6. Final remarks

The results of the theoretical and experimental investigations reported above confirm the possibility of the objective evaluation of the influence of cleft palate, resulting in nasalization, upon the phonetic-acoustical parameters and, particularly, upon the spectral characteristics of oral vowels. This influence is measurable and may be interpreted analytically.

The analysis of the formant-antiformant structure in pathological and postoperative cases of cleft palate gives information on modifications of the acoustical structure of the speech organ due to anatomical disorders. The information is valuable for physicians since it creates new possibilities in objective diagnosis and control of the rehabilitation treatment of cleft palate by means of a spectral analysis of the speech signal.

An essential feature in palatoschisis, of particular interest to the physician, is the extent of anatomical disorders in the palatum and/or velum, expressed in terms of acoustic circuit theory by the shunting effect of the nasal tract. The detailed quantitative investigations of this phenomenon were not the subject of the present work. The results of preliminary experiments proved, however, that objective acoustic methods for phoniatric diagnosis of cleft palate cases may be applied as well to qualitative as to quantitative evaluations of the extent of pathological disorders. Theoretical and experimental research, aimed at these applications, has been recently initiated in the Department of Cybernetics Acoustics of the Institute of Fundamental Technological Research, Polish Academy of Sciences, in close cooperation with the Phoniatric Centre of the Central Clinical Hospital of the Medical Academy in Warsaw.

References

[1] J. BARDACH, *Clefts of the upper lip and of the palatum*, PZWL, Warszawa 1967 (in Polish).

[2] J. W. VAN DEN BERG, J. T. ZANTEMA, P. DOORNENBAL Jr., *On the air resistance and the Bernouilli effect of the human larynx*, JASA, **29**, 5, 626-631 (1957).

[3] G. FANT, *Acoustic theory of speech production*, Mouton and Co., s-Gravenhage 1960.

[4] J. L. FLANAGAN, *Analog measurements of sound radiation from the mouth*, JASA, **32**, 12, 1613-1620 (1960).

[5] J. L. FLANAGAN, *Speech analysis, synthesis and perception*, Springer Verlag, Berlin-Heildelberg-New York, 2nd edition, 1972.

[6] J. KACPROWSKI, *Synthesis of Polish nasal consonants in resonance formant circuits*, Rozprawy Elektrotechniczne, **9**, 3, 439-465 (1963) (in Polish).

[7] P. M. MORSE, *Vibration and sound*, McGraw-Hill Book Co., New York 1948.

[8] 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, **8**, 6a, 70-77 (1974) (in Polish).

Received on 23th September 1975