

Dept. for Speech, Music and Hearing  
**Quarterly Progress and  
Status Report**

**Vocal-tract area and length  
perturbations**

Fant, G.

journal: STL-QPSR  
volume: 16  
number: 4  
year: 1975  
pages: 001-014



**KTH Computer Science  
and Communication**

<http://www.speech.kth.se/qpsr>



## I. SPEECH PRODUCTION

### A. VOCAL-TRACT AREA AND LENGTH PERTURBATIONS

G. Fant

#### Abstract

A circuit-theory approach for deriving formulas for vocal-tract cross-sectional area and length perturbations is outlined. It is demonstrated that the effect of a non-uniform length scaling is proportional to the spatial density of stored reactive energy along the vocal tract. This energy distribution also serves to define the relative cavity-resonance importance of various parts of the tract. The formula is tested in a scaling of vowels with different length factors for the pharynx and the mouth.

### Introduction

This work was initiated in 1966 as a part of a compendium on speech communication for a course at KTH, Fant (1967).

The object of the study has been to derive perturbation formulas from circuit-theory concepts and to extend the theoretical analysis to include non-uniform length perturbation.

### Calculations of resonance frequencies and bandwidths

We shall consider the vocal tract as a transmission line with the distributed inductance and capacitance per unit length

$$\left. \begin{aligned} L(x) &= \rho / A(x) \\ C(x) &= A(x) / \rho c^2 \end{aligned} \right\} \quad (1)$$

where  $A(x)$  is the cross-sectional area at a coordinate  $x$  cm from the glottis. The impedance  $Z(x) = [L(x)/C(x)]^{\frac{1}{2}} = \rho c / A(x)$ . The density  $\rho = 1.2 \cdot 10^{-3}$  g/cm<sup>3</sup> and the velocity of sound  $c = 35300$  cm/s at standard speaking conditions. The distributed series resistance is denoted  $R(x)$  and the shunt conductance  $G(x)$ . Other sources of energy dissipation are the finite glottal resistance  $R_g$ , the resistance of the vocal-tract wall impedance  $R_w$ , and the radiation resistance

$$R_o = \frac{\rho \omega^2 K_s(\omega)}{4 \pi c} \quad (2)$$

The contribution of these dissipative elements to the bandwidth of vocal resonances has been analyzed in detail by Fant (1960) and (1972) and Fant and Pauli (1974). Losses through the wall impedance dominate the bandwidth  $B_1$  of low first formants  $F_1$ . The surface losses through  $R(x)$  and  $G(x)$  determine  $B_2$ ,  $B_3$ , and  $B_4$  when the lipopening is very narrow whilst the radiation resistance  $R_o(\omega)$  is the main determinant of  $B_2$ ,  $B_3$ , and  $B_4$  when the entire vocal tract as well as the lips are open.

From standard circuit theory we can state the following theorem: The poles  $s_n = \sigma_n + j\omega_n$  of a complex network are the same in any transfer function defined by the ratio of any observed current or voltage to a source current or voltage introduced without disturbing the impedance structure. This is simply a consequence of the system determinant being the same. Appropriate sources are series voltages  $e_i$  within a branch

or constant currents feeding into a nodal point. Therefore the poles may be found as the complex frequencies  $s_n$  where the input impedance  $Z_i(s)$ , as seen from any branch, equals zero or

$$\frac{i_i}{e_i} = \frac{1}{Z_i(s)} = \infty \quad (3)$$

In general  $Z_i(s)$  is the sum of the impedances  $Z_l$  looking left and  $Z_r$  looking right

$$Z_i(s_n) = Z_l + Z_r = 0 \quad (4)$$

We shall study the input impedance at the open end of a uniform ideal tube placed inside a spherical baffle of 9 cm radius simulating the head of the speaker. The radiation resistance is intermediate between that of a simple point source and a point source on the surface of an infinite baffle. This relation is carried by the factor  $K_s(\omega)$  in Eq. (2) which varies between 1 and 1.7, see Fant (1960). The radiation inductance is equal to that contained in a lengthening of the tube by  $0.8\sqrt{A_o/\pi}$  cm or

$$L_o = \frac{0.8\sqrt{A_o/\pi} \cdot \rho}{A_o}$$

The condition for finding the poles, Eq. (4), thus implies

$$Z_o \coth\left(\frac{s\ell}{c} + \alpha\ell\right) + R_o(\omega) + sL_o = 0 \quad (6)$$

Within the frequency range of a resonance  $R_o(\omega)$  is approximately a constant. The attenuation constant  $\alpha = \frac{R}{2Z_o} + \frac{G \cdot Z_o}{2}$  includes the surface losses.

After inserting  $s = \sigma + j\omega$  and expansion of  $\coth\left(\frac{s\ell}{c} + \alpha\ell\right)$  we equate the real and the imaginary parts of the equation separately to zero.

Since the internal attenuation constant  $\alpha \ll 1$  and  $\alpha_o = R_o/Z_o < 0.4$ , we may introduce approximations which reduce Eq. (6) to

$$\operatorname{tg} \frac{\omega_n \ell}{c} = \frac{Z_o}{\omega_n L_o} \quad (7)$$

as in the loss-less case and

$$\sigma = \frac{-[R_o(\omega) + \alpha\ell \cdot Z_o]}{L_o + Z_o \ell/c} \quad (8)$$

After introducing the effective length of the tube

$$\left. \begin{aligned} l_e &= l + l_o \\ l_o &= 0.8\sqrt{A_o/\pi} \end{aligned} \right\} \quad (9)$$

Eq. (8) may be rewritten

$$\sigma_n = \frac{-R_o(\omega_n)}{Z_o \cdot l_e/c} - \alpha \cdot \frac{c \cdot l}{l_e} \quad (10)$$

or

$$B_n = -\frac{\sigma_n}{\pi} = \frac{R_o(\omega_n)}{Z_o} \cdot \frac{c}{\pi \cdot l_e} + \frac{\alpha \cdot c \cdot l}{\pi \cdot l_e} \quad (11)$$

which we denote

$$B_n = B_{on} + B_{in} \quad (12)$$

In the Appendix of Fant (1960) it is shown that resonance frequencies and bandwidths calculated according to Eqs. (7) and (11) agree well with measurements on a physical model (wooden sphere as a baffle for a brass tube of 16.4 cm length and 8 cm<sup>2</sup> area,  $c = 34.400$  cm/s at 20°C). These data are summarized in the following tabulation:

TABLE I-A-I.

	$F_n$		surface losses		radiation		total bandwidth	
			$B_{in}$	$B_{on}$	$R_o/Z_o$	$K_S(\omega)$	$B_n$	$B_n$
	Calc.	Meas.	Calc.	Calc.	Calc.	Calc.	Calc.	Meas.
	Hz	Hz	Hz	Hz			Hz	Hz
F1	486	485	4.5	3.4	0.0055	1.1	7.9	8
F2	1459	1459	7.9	43.6	0.07	1.6	51.5	44
F3	2445	2434	10.2	133	0.20	1.7	143	128
F4	3444	3442	12.1	225	0.34	1.45	237	228

This experiment (carried out in cooperation with A. Møller in 1957) verifies the assumptions of circuit constants as well as the theoretical approach. The usefulness of the concept of the effective length  $l_e = l + l_o$  was demonstrated in the bandwidth calculations. Instead of the more exact Eq. (7) we may calculate resonance frequencies of the single tube from

$$\cot \frac{\omega_n \cdot l_e}{c} = 0 \quad (13)$$

The errors in the first four formants are 0, -2, -16, -43 Hz, respectively. With smaller lipopenings the errors become smaller.

The distribution of volume velocity and pressure inside the tract is not affected by inclusion of the loss elements. This has been demonstrated by Mrayati and Carré (1975) in a report where the calculations of Fant and Pauli (1974) have been applied to French vowels. The effect is most noticeable in open vowels where the radiation impedance dominates the loss elements. Also the spatial distributions of pressure in the complete LEA-model for the Russian vowels agree well with those of the loss-less distributions calculated by Fant and Pauli (1974). In the general case of an arbitrary vocal-tract configuration the bandwidths may be calculated from the ratio of dissipated and stored energies per cycle of a mode, as demonstrated by Fant and Pauli (1974) or from the interpolation of a complex system determinant, Fant (1970).

#### Perturbation theory

The approach followed here is to study the separate effects of changes,  $\Delta L(x)$  and  $\Delta C(x)$ , in the distributed elements  $L(x)$  and  $C(x)$ , and to relate the observed frequency shifts to equivalent changes,  $\Delta A(x)$  and  $\Delta(x)$  in cross-sectional area and in unit length of a section of the transmission line. Two different methods leading to the same result have been developed. One is concerned with the impedance condition for resonance as outlined in the previous section, the other starts out from the criterion of kinetic energy of a mode equaling the potential energy of the mode.

As already discussed in the previous section, Eq. (4), the poles are found as the complex frequencies providing zero branch impedance, as seen from the element  $\Delta L(x)$ .

$$Z_r(x, s) + Z_\ell(x, s) + s \cdot \Delta L(x) = 0 \quad (14)$$

Neglecting losses (as proved to be possible in the open tube case) the resonance frequencies  $\omega_{n1}$  before the perturbation satisfy

$$X(x, \omega_{n1}) = X_r(x, \omega_{n1}) + X_\ell(x, \omega_{n1}) = 0 \quad (15)$$

Since any reactance increases with frequency, the insertion of  $\Delta L(x)$  causes a negative frequency shift  $\Delta \omega_{n1} = \omega_{n2} - \omega_{n1}$  of the resonance mode  $\omega_n$ .

$$\Delta \omega_{n1} \frac{dX(x, \omega)}{d\omega} + (\omega_{n1} + \Delta \omega_{n1}) \Delta L(x) = 0 \quad (16)$$

and

$$\frac{\Delta \omega_{n1}}{\omega_{n1}} = \frac{-\Delta L(x)}{X'(\omega_{n1}) + \Delta L(x)} \quad (17)$$

Similarly, after insertion of  $\Delta C(x)$ , there is a shift  $\Delta \omega_{n2} = \omega_{n3} - \omega_{n2}$  determined by:

$$\Delta \omega_{n2} \frac{dB(x, \omega)}{d\omega} + (\omega_{n2} + \Delta \omega_{n2}) \Delta C(x) = 0 \quad (18)$$

where the susceptance

$$B = \frac{-1}{X_r} + \frac{-1}{X_l} \quad (19)$$

and

$$\frac{\Delta \omega_{n2}}{\omega_{n2}} = \frac{-\Delta C(x)}{B'(\omega_{n2}) + \Delta C(x)} \quad (20)$$

From Eqs. (19) and (15) we find

$$B'(x, \omega) = \frac{X'(x, \omega)}{X_r^2(x, \omega)} = \frac{X'(x, \omega) \cdot U^2(x, \omega)}{P^2(x, \omega)} \quad (21)$$

It is now convenient to convert  $X'$  to an energy function. The power input to a two-terminal network with input impedance  $Z_i = P_i/U_i$  is  $Z_i \cdot U_i^2$ . If all branches  $k$  contain reactances,  $\omega L_k$  and/or  $\frac{-1}{\omega C_k}$ , the sum of the reactive power stored in the reactances equals the reactive power input to the terminal

$$X_i \cdot U_i^2 = \sum \left[ \omega L_k - \frac{1}{\omega C_k} \right] U_k^2 \quad (22)$$

After differentiation

$$\frac{1}{2} \cdot \frac{dX_i(\omega)}{d\omega} U_i^2 = \sum \frac{1}{2} U_k^2 L_k + \sum \frac{1}{2} U_k^2 \frac{1}{\omega^2 C_k} = E_K + E_P \quad (23)$$

Here  $E_T = E_K + E_P$  is the sum of the total kinetic  $E_K$  and the total potential energy  $E_P$ . From Eqs. (17) and (23),

$$\frac{\Delta \omega_{n1}}{\omega_{n1}} = \frac{-\Delta L(x) \cdot U^2(x) \frac{1}{2}}{E_T + \Delta L(x) U^2(x) \frac{1}{2}} = \frac{-\Delta E_K(x)}{E_T + \Delta E_K(x)} \quad (24)$$

$$\frac{\Delta \omega_{n1} + \omega_{n1}}{\omega_{n1}} = \frac{\omega_{n2}}{\omega_{n1}} = \frac{E_T}{E_T + \Delta E_K(x)} \quad (25)$$

where  $\Delta E_K(x)$  is the increase of kinetic energy associated with the element  $\Delta L(x)$  and a constant current  $U(x)$ .

Similarly from Eqs. (20), (21), and (23)

$$\frac{\Delta \omega_{n2}}{\omega_{n2}} = \frac{-\frac{1}{2} P^2(x) \Delta C(x)}{E_T + \frac{1}{2} P^2(x) \Delta C(x)} = \frac{-\Delta E_P(x)}{E_T + \Delta E_P(x)} \quad (26)$$

It is here understood that  $P(x)$  pertains to the pressure before insertions of  $\Delta L(x)$  and that  $E_T$  thus is the same as in Eqs. (24) and (26).

$$\frac{\omega_{n3}}{\omega_{n2}} = \frac{E_T}{E_T + \Delta E_P(x)} \quad (27)$$

Combining Eqs. (25) and (27) we get a total frequency shift which is approximately

$$\frac{\omega_{n3}}{\omega_{n1}} = \frac{E_T}{E_T + \Delta E_P(x) + \Delta E_K(x)} \quad (28)$$

or

$$\frac{\Delta \omega}{\omega_{n1}} = \frac{\omega_{n3} - \omega_{n1}}{\omega_{n1}} = \frac{-[\Delta E_K(x) + \Delta E_P(x)]}{E_T + \Delta E_K(x) + \Delta E_P(x)} \quad (29)$$

We may now study the effect of a perturbation in  $A(x)$ .

For small perturbations  $\Delta A(x)/A(x) \ll 1$ , we can write

$$\left. \begin{aligned} \Delta E_K(x) &= \frac{\Delta L(x)}{L(x)} \cdot E_K(x) = \frac{-\Delta A(x)}{A(x)} E_K(x) \\ \Delta E_P(x) &= \frac{\Delta A(x)}{A(x)} \cdot E_P(x) \end{aligned} \right\} \quad (30)$$

The combined effect of distributed perturbations is then approximately

$$\frac{\Delta \omega}{\omega} = \frac{\Sigma [E_K(x) - E_P(x)] \frac{\Delta A(x)}{A(x)}}{\Sigma E_K(x) + \Sigma E_P(x)} \quad (31)$$

as stated by Fant (1967), Fant and Pauli (1974). To derive the effects of lengths perturbations  $\Delta(x)$ , we express the change in kinetic energy per unit length

$$\left. \begin{aligned} \Delta E_K(x) &= \frac{\Delta(x)}{1} \cdot E_K(x) \\ \Delta E_P(x) &= \frac{\Delta(x)}{1} \cdot E_P(x) \end{aligned} \right\} \quad (32) \quad (32)$$

Retaining the denominator of Eq.(29) for increased accuracy and denoting  $E_K(x) + E_P(x) = E_T(x)$ , we find

$$\frac{\Delta \omega}{\omega} = \frac{-\Sigma \Delta(x) E_T(x)}{\Sigma [1 + \Delta(x)] E_T(x)} \quad (33)$$

The alternative approach is to start out from the pressure-velocity equations for the transmission line analog

$$\left. \begin{aligned} \frac{dP(x)}{dx} &= -j\omega U(x) \cdot L(x) \\ \frac{dU(x)}{dx} &= -j\omega P(x) \cdot C(x) \end{aligned} \right\} \quad (34)$$

A perturbation  $\Delta L(x)$  introduces an excess of kinetic energy  $\Delta E_K(x) = \frac{1}{2} U^2(x) \Delta L(x)$ , which is compensated for by a frequency lowering that increases the potential energy and/or decreases the kinetic energy in other parts of the line. At resonance, before and after the perturbation the kinetic energy equals the potential energy but we are free to chose any relation between the actual energy levels before and after the perturbation. Assume  $U(x)$  and thus  $dU/dx$  approximately the same after perturbation.

	Kinetic energy	Potential energy	
Before pertur- bation	$E_K$	$= \Sigma \left[ \left( \frac{dU}{dx} \right)^2 \frac{1}{\omega_1^2 C(x)} \right]$	} (35)
After $\Delta L(x)$ pertur- bation	$E_K + \Delta E_K(x)$	$= \Sigma \left[ \left( \frac{dU}{dx} \right)^2 \frac{1}{\omega_2^2 C(x)} \right]$	

or

$$\frac{\omega_2^2}{\omega_1^2} = \frac{E_K}{E_K + \Delta E_K(x)} \quad (36)$$

and

$$\frac{\omega_2}{\omega_1} = \frac{E_K}{E_K + \frac{1}{2} \Delta E_K(x)} = \frac{E_T}{E_T + \Delta E_K(x)} \quad (37)$$

which is identical to Eq. (25). Similarly, a perturbation  $\Delta C(x)$  assuming  $P(x)$  and  $dP(x)/dx$  to be approximately constant is associated with the following energy states

Kinetic energy	Potential energy	
Before $\Delta C(x)$ pertur- bation	$\Sigma \left[ \left( \frac{dP}{dx} \right)^2 \frac{1}{\omega_2^2 L(x)} \right]$	$E_P$
After $\Delta C(x)$ pertur- bation	$\Sigma \left[ \left( \frac{dP}{dx} \right)^2 \frac{1}{\omega_3^2 L(x)} \right]$	$E_P + \Delta E_P(x)$

$\left. \vphantom{\begin{matrix} \text{Before} \\ \text{After} \end{matrix}} \right\} \quad (38)$

or

$$\left( \frac{\omega_3}{\omega_2} \right)^2 = \frac{E_P}{E_P + \Delta E_P(x)} \quad (39)$$

Since  $E_P = \frac{1}{2} E_T$  we may approximate

$$\left( \frac{\omega_3}{\omega_2} \right) = \frac{E_T}{E_T + \Delta E_P(x)} \quad (40)$$

as in Eq. (27). For distributed perturbations, Eqs. (36) and (39) combine to

$$\left( \frac{\omega_3}{\omega_1} \right)^2 = \frac{1}{(1 + \Sigma [\Delta E_K(x)/E_K(x)]) (1 + \Sigma [\Delta E_P(x)/E_P(x)])} \quad (41)$$

### Discussion and experimental validation

The relation between energy and frequency shifts was discussed by Schroeder (1967) with reference to perturbations in the uniform tube. These relations, referred to as the Ehrenfest formula, are approximately valid for an arbitrary shaped vocal tract, Fant (1967). Instead of the simpler formula  $\Delta \omega / \omega = -\Delta E / E$  we retain a higher degree of accuracy by the relation  $\Delta \omega / \omega = -\Delta E / (E + \Delta E)$ , as pointed out by Schroeder. We shall see how this is validated by a uniform length perturbation, i. e. a simple scale factor change in length dimension. Instead of  $\Delta(x)$  in Eq. (33), we introduce the constant  $\Delta$ , which we relate to a new parameter

$$k = - \frac{\Delta}{1 + \Delta} \quad (42)$$

Eq. (33) thus reduces to

$$\frac{\Delta \omega}{\omega} = k \quad (43)$$

As an example, if a single tube resonator is doubled in length, i. e.  $\Delta = 1$ , the frequency shift is  $\Delta\omega/\omega = -0.5$ , which also holds for a general shaped vocal tract. Eq. (43) implies that all formants are shifted by  $k\%$ , when an overall scale factor in length of  $\Delta = -k/(1+k)$  is applied. As verified numerically we can approximate Eq. (33) very well with the expression

$$\frac{\Delta\omega}{\omega} = \frac{-\sum \left[ \frac{\Delta(x)}{1+\Delta(x)} E_T(x) \right]}{\sum E_T(x)} = \frac{\sum [k(x)E_T(x)]}{\sum E_T(x)} \quad (44)$$

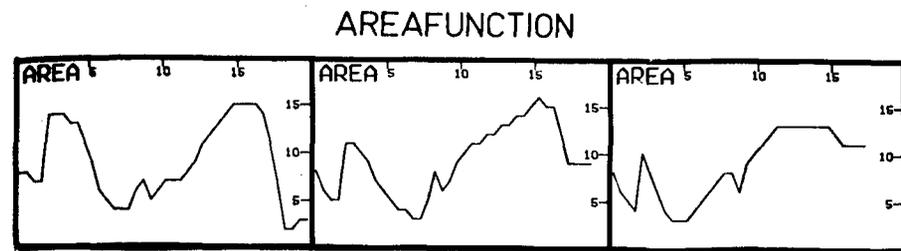
which means that the relative frequency shift  $\Delta\omega/\omega$  is the energy weighted average of the frequency shift factors  $k(x)$  applied to the separate parts of the vocal tract.

Uniform scaling in vocal-tract area,  $\Delta A(x)/A(x) = \text{constant}$ , should leave all resonance frequencies intact. This may be derived from perturbation formula, Eq. (31), since  $\sum E_K(x) = \sum E_P(x)$ . For greater values of  $\Delta A(x)/A(x)$  we have to write  $\Delta L/L = -\Delta A/(A+\Delta A)$  instead of  $-\Delta A/A$  and it is only Eq. (41) that then stands up to the test of  $\Delta\omega/\omega = 0$ .

We are now in a position to make a more precise statement about the uniform length scaling. The inverse proportionality of total length and resonance frequencies holds exactly only if the end correction  $0.8\sqrt{A_0/\pi}$  is scaled by the same factor as length dimensions. This condition would hold for a simultaneous scaling of length and vocal-tract area by the same linear factor.

It was suggested by Fant (1960) and again by Fant and Pauli (1974) that the sum of the spatial kinetic and potential energy densities, i. e.  $E_T(x)$ , is a suitable parameter for describing the "cavity-resonance" dependency. This is intuitively supported by the fact of  $E_T(x)$  being constant along a uniform tube resonator, all parts contributing the same to the tuning of the mode.

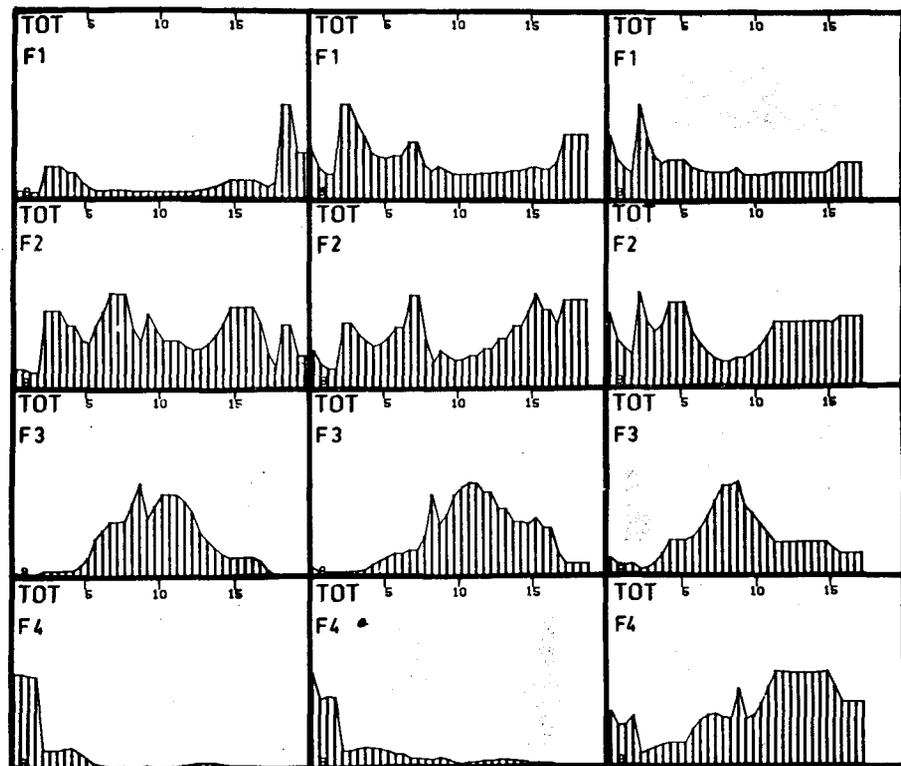
In addition, we have now proved that the energy density is a measure of the sensitivity of the tract to local expansions or contractions of the length dimension. In Fig. I-A-1, pertaining to the six Russian vowels studied by Fant (1960) and processed by Fant and Pauli (1974), this parameter  $E_T(x)$  is denoted  $TOT = KIN+POT$  whilst the area perturbation sensitivity  $LAG = KIN-POT$  of the same vowel is shown in Fig. I-A-2.



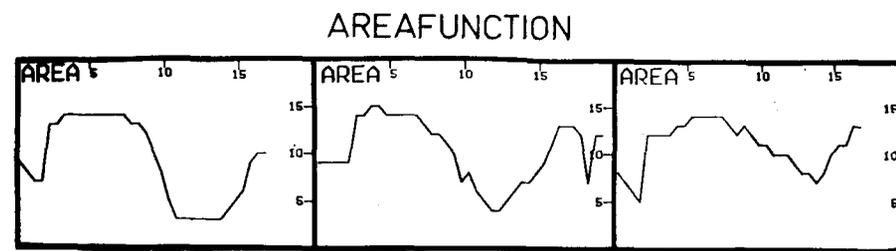
[u]

[o]

[a]



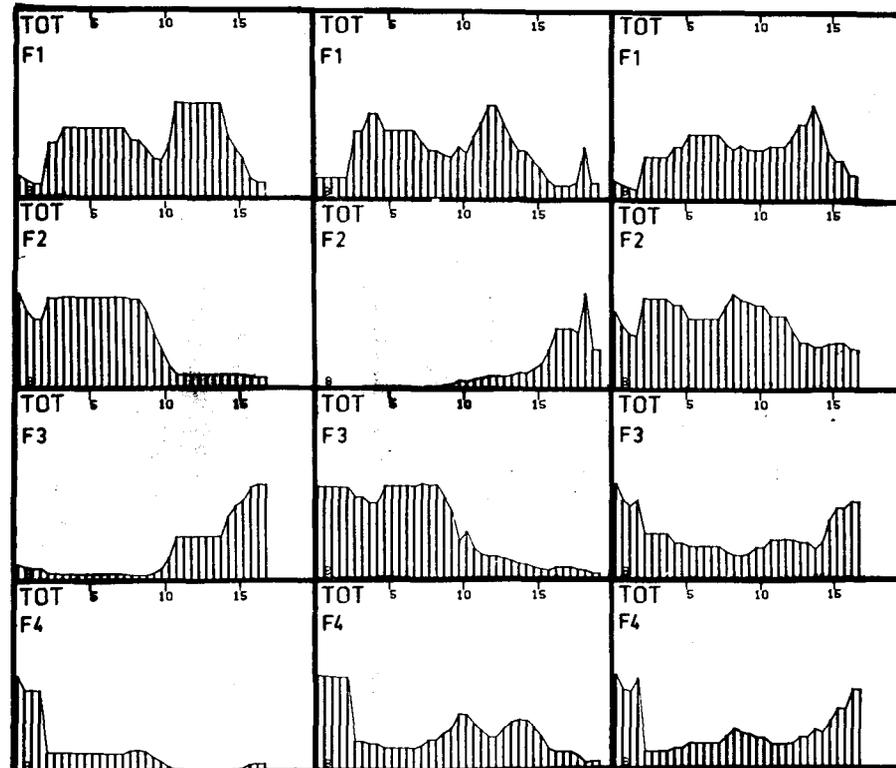
KIN+ POT ENERGY DENSITY



[i]

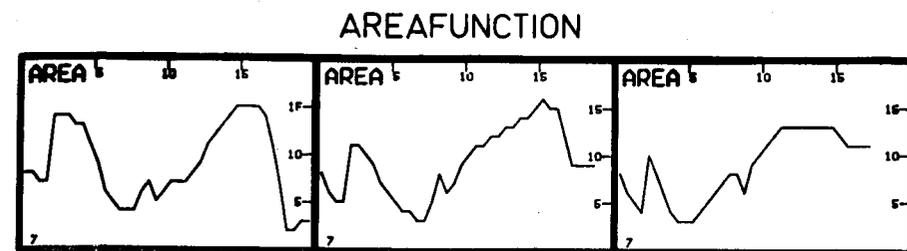
[ɨ]

[e]

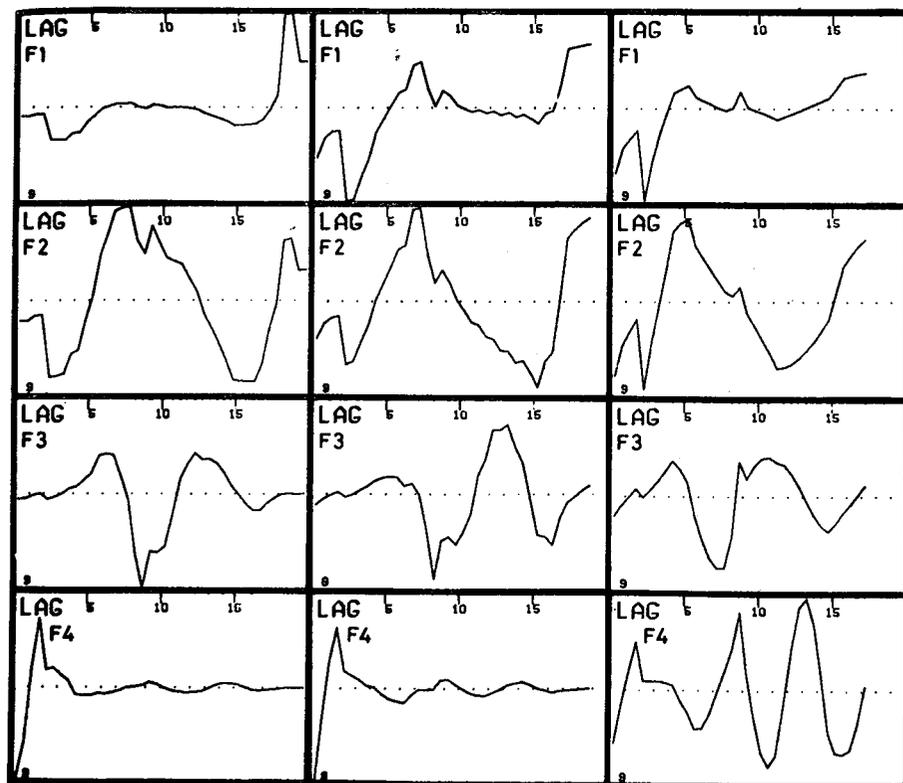


KIN + POT ENERGY DENSITY

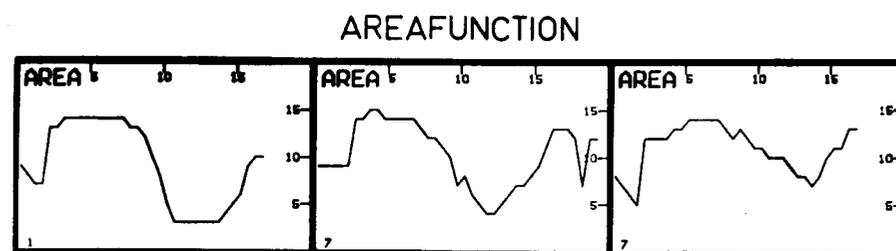
Fig. I-A-1. Spatial distribution of kinetic plus potential energy for the six Russian vowels of Fant (1960). From Fant and Pauli (1974). These graphs are useful for predicting the effects of length perturbations.



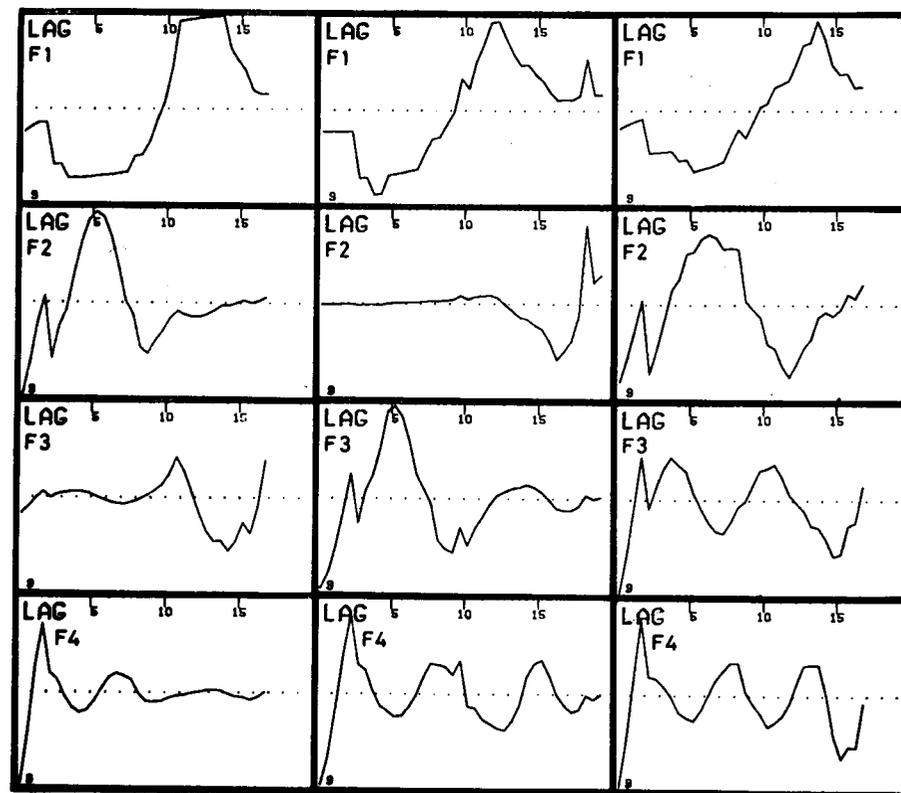
[u] [o] [a]



LAG=KIN-POT ENERGY DENSITY



[i] [ɨ] [e]



LAG=KIN-POT ENERGY DENSITY

Fig. I-A-2. Spatial distribution of kinetic energy KIN minus the potential energy POT for the six Russian vowels. This parameter, denoted LAG, the "Lagrangian", displays the sensitivity to local area changes.

As expected, in the regions of maximum KIN+POT the parameter KIN-POT shows its largest oscillations. As has been pointed out many times before, e. g. by Fant (1960) and Fant and Pauli (1974), we observe the front part of the tract affiliation of  $F_2$  of [ɪ] and  $F_3$  of [i] and the back-part affiliation of  $F_3$  of [ɪ] and  $F_2$  of [i]. There is also a peak of prominence at the lipopening of  $F_1$  of [u] and at the larynx tube for  $F_4$  of all vowels. We can also observe the distribution of F3-energy of [u], [o], and [a] in the middle part of the tract with a tendency of the F3-[o]-energy to have its spatial peak in the mouth part of the tract. These observations conform well with the results discussed by Fant (1975) in the previous issue of the STL-QPSR. The specific sensitivities of vowel formants to different scalings in the mouth and in the pharynx follow the energy distributions. However, the main tendency as observed by Nordström (1975) is that of uniform shift of resonance frequencies in spite of separate scaling factors being applied to the mouth and the pharynx. This average tendency conforms with the general impression of Fig. I-A-1 that in most vowels the energy of any mode is substantially spread out over the entire tract. In a more detailed view, we observe the non-uniform energy distributions discussed above.

We shall now test the perturbation formula for separate scale factors in the mouth and in the pharynx. Eq. (44) becomes

$$\frac{\Delta \omega}{\omega} = \frac{k_1 \cdot E_1 + k_2 \cdot E_2}{E_1 + E_2} \quad (45)$$

where  $E_1$  is the total energy in the pharynx (from the glottis to the uvular region) and  $E_2$  is the total energy in the mouth (from the uvular region to the point of radiation at the lips). If the pharynx is scaled to  $\Delta_1 \cdot 100$  per cent increase in length, the associated frequency factor  $k_1$  is defined as  $-\Delta_1 \cdot 100 / (1 + \Delta_1)$  per cent.

Results of perturbation of the [i]-tract by use of Eq. (45) are shown in the following tabulation together with exact values calculated from the Liljencrants-Fant (1975) program for converting area functions to resonance frequencies.

TABLE I-A-II. Length perturbation of the [ i ]-vowel.

Scaling	$k_m$ %	$k_p$ %	k %	$\Delta/\omega$ Exact %		$\Delta\omega/\omega$ Perturb. %	
				$F_2$	$F_3$	$F_2$	$F_3$
1. $\Delta_m = -0.05$ $\Delta_p = -0.15$	5.3	17.6	11.6	14.5	7.5	15.5	7.7
2. $\Delta_m = -0.23$ $\Delta_p = -0.36$	30	56	42	48	35.5	51	35
3. = 2. with m-p divi- sion shift- ed x = -1.5 cm	30	56	42	45	34		35
4. = 3. but no scaling of end cor- rection				45.5	32		
5. $\Delta_m = -0.11$ $\Delta_p = -0.18$	12.4	22	17.5	20	14	20	14.5

In scalings 1.-4. the end correction is scaled with the same factor as the mouth. The constant end correction in 4. accounts for a larger radiation inductance which counteracts the mouth cavity length reduction and thus the relative  $F_3$ -increase. The conditions of scaling 5. provide  $F_2$ - and  $F_3$ -shifts which are typical of females compared to males. Here the mouth cavity is shortened 11 %, the pharynx by 18 %. The perturbation formula provides almost the same formant shifts as the exact calculations, 20 % in  $F_2$  and 14 % in  $F_3$ . If the energy of these resonances had been confined to the pharynx only and the mouth only, the shifts had been 22 % and 12.4 %, respectively. In the six language study of Fant (1975) the average female-male differences were 21 % in  $F_2$  and 13 % in  $F_3$ . The particular convention for deciding the boundary between the pharynx and the larynx has some influence on the  $F_2$ -shift observed. The boundary point selected by Nordström (1975) was 1.5 cm below that of mine, which accounts for a 3 % smaller shift in  $F_2$  and 1.5 % smaller shift in  $F_3$ , compare scalings 3. and 2. in Table I-A-II.

One shortcoming of the energy distributions in Figs. I-A-1 and I-A-2 is that they do not include the kinetic energy stored in the radiation reactance. However, it has been extrapolated and made use of in the present perturbation calculations. It would be logical to extend the x-coordinate of the energy function by the radiation inductance end-correction length and to confine the kinetic energy of radiation  $\frac{1}{2} L_o U_o^2$  to this interval instead of lumping it all to the x=0 coordinate, as done by Mrayati and Carré (1975).

Fig. I-A-2 serves as a guide for assessing how sensitive any mode is to local area changes at any coordinate. These curves should not be confused with the distributed perturbations affecting one resonance only, as derived for instance by Heinz (1967).

A few characteristics can be mentioned. In all vowels  $F_1$  is raised by a contraction of the pharynx. The expansion of the mouth cavity is more effective in rising  $F_1$  of front vowels than back vowels. An exception is the apparent influence of a lipopening in rising  $F_1$  of [u] and [o]. In [o] and [a] it is a narrow region a few cm above the larynx where a contraction is maximally effective in rising  $F_1$ . This might be a factor of relevance to the large female-male difference in  $F_1$  of maximally open vowels.

$F_2$  of [u], [o], and [a] are raised by an expansion of the middle or upper part of the pharynx or by a contraction of the mouth cavity. An expansion of the middle part of the pharynx of [i] or of the upper part of the pharynx of [e], raises  $F_2$ .  $F_2$  of [ɨ] is almost insensitive to pharyngeal perturbations. In all the back vowels  $F_3$  is raised by a contraction in the uvular region.

#### References

- FANT, G. (1960): Acoustic Theory of Speech Production, Mouton, 's-Gravenhage (2nd edition 1970).
- FANT, G. (1967): "Kompendium i talöverföring", Del 1, KTH, Dept. of Speech Communication.
- FANT, G. (1972): "Vocal-tract wall effects, losses, and resonance bandwidths", STL-QPSR 2-3/1972, pp. 28-52.
- FANT, G. (1975): "Non-uniform vowel normalization", STL-QPSR 2-3/1975, pp. 1-19.
- FANT, G. and PAULI, S. (1974): "Spatial characteristics of vocal-tract resonance modes", paper presented at the Speech Communication Seminar, Stockholm, Aug. 1974; to be publ. in Speech Communication, Vol. 2 by Almqvist & Wiksell, Stockholm.

- HEINZ, J. M. (1967): "Perturbation functions for the determination of vocal-tract area functions from vocal-tract eigenvalues", STL-QPSR 1/1967, pp. 1-14.
- LILJENCRAANTS, J. and FANT, G. (1975): "Computer program for VT-resonance frequency calculations", this issue of STL-QPSR, pp. 15-20.
- MRAYATI, M. and CARRE, R. (1975): "Acoustic aspects of French nasal vowels", paper presented to the 89th Meeting of the Acoustical Society of America, Austin, Texas.
- NORDSTRÖM, P. -E. (1975): "Attempts to simulate female and infant vocal tracts from male area functions", STL-QPSR 2-3/1975, pp. 20-33.
- SCHROEDER, M. R. (1967): "Determination of the geometry of the human vocal tract by acoustic measurements", J. Acoust. Soc. Am. 41, pp. 1002-1010.