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

Notes on vocal tract computation

Badin, P. and Fant, G.

journal: STL-QPSR
volume: 25
number: 2-3
year: 1984
pages: 053-108

KTH Computer Science and Communication

C. NOTES ON VOCAL TRACT COMPUTATION

Pierre Badin* and Gunnar Fant

Abstract

In the general frame of vocal tract computation in the frequency domain, we describe experiments aiming at the evaluation of different methods to handle the boundary conditions and the losses, that is, the radiation load, the viscous and thermal losses. the wall vibrations, the glottal and subglottal impedances, and the constriction resistances. We also discuss a method to determine poles and zeros for any transfer function in the tract. We give data for the Russian vowels (FANT, 1960) in different conditions, including diver's speech, and show an example of application to the study of a constrictive consonant.

Introduction

Vocal tract computations in the frequency domain is nowadays well established (FANT, 1960; PORTNOFF, 1973; MRAYATI, 1976; ATAL et al., 1978; WAKITA & FANT, 1978; FANT, 1985). Thus, we will not enter into detailed derivations of the classical equations and electrical network representations of the acoustic propagation in the vocal tract. Our aim, in this study, was to compare different methods to handle the boundary conditions and the losses in the tract - that is, the radiation load, the viscous and thermal losses, the wall vibrations, the glottal and subglottal impedances, and the constriction resistances -, in order to achieve a better accuracy and computational efficiency, and to evaluate the consequences of certain simplifications and alternative transfer function algorithms. We mainly used the Russian vowels (FANT, 1960) for these evaluations, and included some calculations on helium speech. We also illustrate a simulation of constrictive consonants.

1. Vocal tract computation: general frame

In this part, we outline the purpose and function of the vocal tract computations: what kind of input data is to be processed, and what kind of data can be derived. Then we describe the computational principles. Comments are made on methods for pole/zero determinations.

1.1 INPUT AND OUTPUT DATA FOR THE SIMULATION

The general purpose of the system is to handle any vocal tract configuration, and to derive the frequency transfer functions between a pressure or volume velocity source in the vocal tract and the output flow at the lips or radiated pressure, and the complex poles and zeros of these functions.

---

* Guest researcher at KTH, March 1983-November 1984, granted from the Institut National de Recherche en Informatique et Automatique, Paris (France), and from KTH, Dept. of Speech Communication & Music Acoustics.

The input data consist of two sets of items:

(1) Area function (the vocal tract is decomposed into a series of finite length-tubes and -horns); location of sources, wall impedances and constriction resistance can be specified also;

(2) Boundary conditionsand other features(type of losses, radiation impedance, etc...); state of the glottis.

The area function values are put in a file; the boundary features can be entered by means of the keyboard (see Fig. 16b).

The output data consist of the frequency transfer function between the source and the output flow at the mouth, or the radiated pressure, and of its complex poles and zeros (see Figs. 16c and 16d).

1.2 VOCAL TRACT CONFIGURATION AND COMPONENTS

1.2.1 Vocal tract configuration

The basic configuration (see Fig. 1) is a single transmission line for the subglottal system, a glottal impedance, and a subglottal network. Cavity wall loading is introduced by distributed or lumped impedances shunting the line (see Fig. 11). A nasal branch is under development. A tentative model of the nasal system including two shunting nasal sinuses has given promising results. It will be described in a separate report. We are eventually aiming at a complete modular system including the mixing in of oral and nasal outputs and the externally radiated sounds from the walls of the tract. We do not intend to review the entire acoustic theory of speech production, but we have attempted to test and to clarify some of the formulas and computational procedures.

The subglottal system is represented by a three cell Foster network (ANANTHAPADMANABHA & FANT, 1982). The glottis is represented by a resistor and an inductor (see Section 2.4). The vocal tract itself is composed of a series of electrical quadripoles corresponding to finite length-cylindrical tubes, -conical horns, parallel impedances, and series impedances. A pressure source can be inserted between any two quadripoles. The basic frequency function to be computed is either the transconductance between the mouth output flow I_0 and the pressure source E_g or E_c, or the ratio between I_0 and I_q, I_q being the input current to the vocal tract if the subglottal system is supposed to have an infinite impedance.

1.2.2 The electrical quadripoles

1.2.2.1 Cylindrical tube sections

The derivation which leads from the basic equations of the acoustic waves propagation in a cylindrical tube to an equivalent electrical quadripole representation is well known and has been extensively treated in the literature (FANT, 1960; FLANAGAN, 1972; MRAYATI, 1976). In our model, we make use of the network of Fig. 2, taken from FANT (1960).

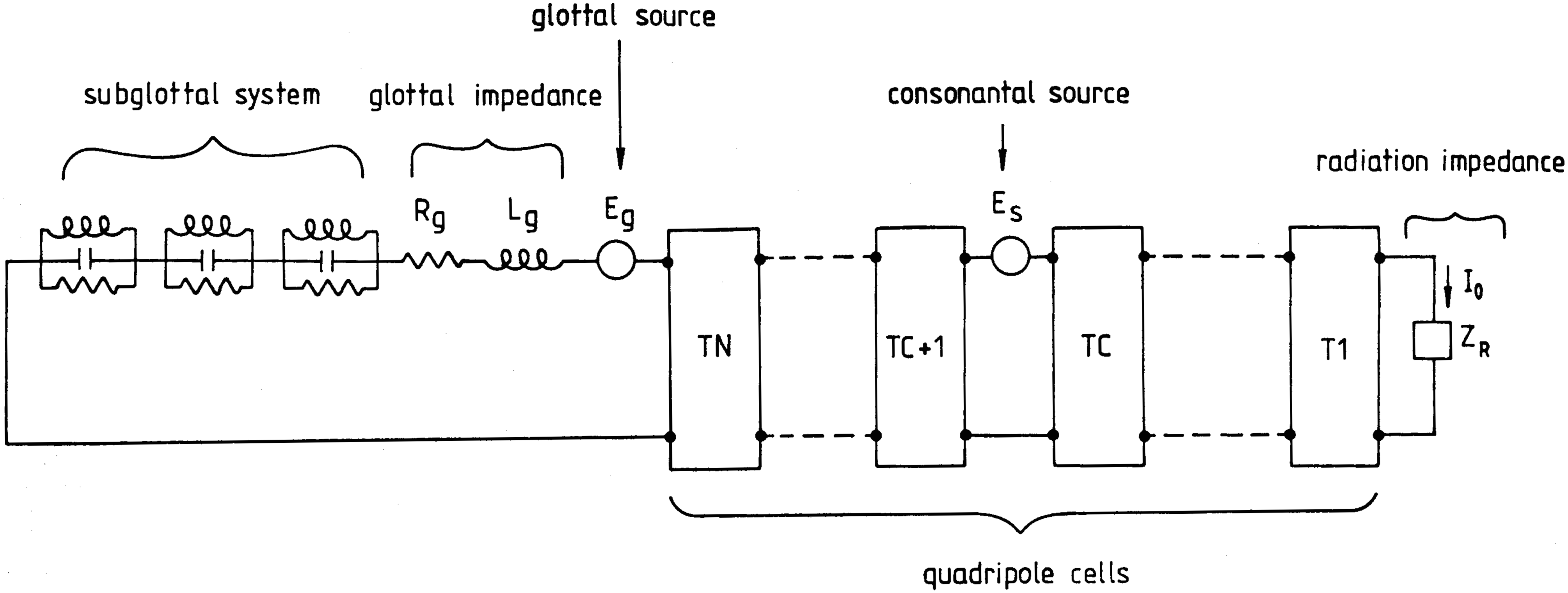
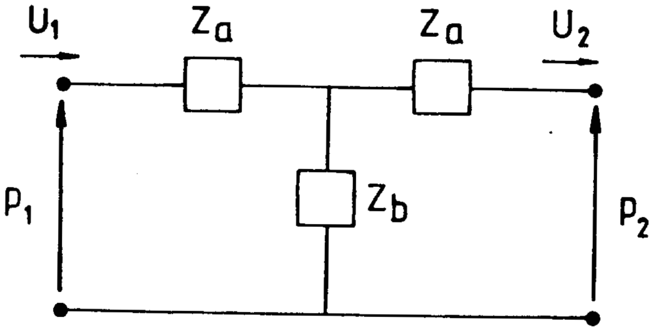


Fig. 1. General structure of the vocal tract simulation.



$$Z_a = Z_c \cdot \tanh(\Gamma \cdot l/2)$$
$$Z_b = Z_c / \sinh(\Gamma \cdot l)$$

$$\begin{cases} \Gamma = \alpha_R + \alpha_G + j\beta \\ Z_c = Z_o \cdot \;\; 1 - j\,\dfrac{\alpha_R - \alpha_G}{\beta} \end{cases}$$

$$\begin{cases} \beta = \dfrac{\omega}{c} & Z_o = \dfrac{\rho c}{A} \\ \alpha_R = \dfrac{R \cdot A}{2\rho c} & \alpha_G = \dfrac{G\rho c}{2A} \end{cases}$$

(R and G are defined in § 2.2.1)

Fig. 2. Network representation of cylindrical tubes.

1.2.2.2 Horn sections

Cylindrical sections may conveniently be combined with horn-shaped modules for modeling a vocal tract area function. Wave equations for acoustic propagation in horns have been derived by MORSE (1976) and equivalent circuits are given by FANT (1960). A corrected version of the FANT (1960) models is shown in Fig. 3. It should be observed that these hold for the case of area independent losses only. A technique for overcoming this limitation is described in Section 2.7.

1.2.2.3 Parallel and series impedances

It is possible to insert either a parallel or a series impedance between two quadripoles in the chain (see Fig. 4). Impedances paralleling the line are used to simulate wall effects (see Section 2.3.2). This technique also applies to the representation of nasal sinuses. Series impedances are used to represent additional losses at a constriction (see Section 2.6) or internal end corrections of radiation inductance from a narrow tube abruptly terminated in a wider tube, see FANT (1960, p. 36). This latter term has not been included in the examples of the present study, unless it is explicitely mentioned. Parallel and series impedances also enter the horn models.

1.2.3 Computation principles

The general purpose is to derive the frequency properties of the network, that is, input impedances at certain locations in the system and transfer functions between a pressure or a flow at a given location and another pressure or flow at another location.

Different techniques can be used in order to compute this type of functions. One well known is the matrix technique used by WAKITA & FANT (1978) or ATAL et al. (1978). For a given frequency, each electrical quadripole is represented by a matrix which gives the input flow and pressure as a linear combination of the output flow and pressure. One just needs to establish the product of all these matrices corresponding to the vocal tract to obtain the transfer function.

An alternative is the method used by FANT (1960, p.38) and by LILJENCRANTS & FANT (1975). It operates on hyperbolic finite length acoustic tube analogs, which ensures greater accuracy and flexibility than the lumped element modular approximation adopted by WAKITA & FANT (1978). The input impedance of the cell of Fig. 5 may be written

$$Z_{i,n} = Z_n \cdot \mathrm{tgh}(\theta_n + \theta_{i,n-1}), \tag{1}$$

with

$$\theta_{i,n-1} = \mathrm{artgh}(Z_{i,n-1}/Z_n). \tag{2}$$

Thus, by reccurence, the input impedance of the cell n is a function of the input impedance to cell n-1.

It is a classical theorem of electrical networks theory that the input impedance at any point in the network has the same poles as any transfer function between two points. It is thus possible to derive the poles from an input impedance. This technique is mainly of interest for loss-free treatment of vowels. A disadvantage is that pole determinations may become contaminated when zeros appear close to poles.

We shall now describe a new computationnally more efficient technique which leads itself to the handling of zero free transfer functions such as the overall volume velocity transfer I_0/I_g. From Fig. 2, we may derive a system of equations, for the cell number n, with flow and pressure inputs U_n and P_n:

$$\begin{cases} P_n = (Z_n \coth\theta_n) \cdot U_n - \dfrac{Z_n}{\sinh\theta_n} \cdot U_{n-1} & \text{(3a)} \\[2ex] P_{n-1} = \dfrac{Z_n}{\sinh\theta_n} \cdot U_n - (Z_n \coth\theta_n) \cdot U_{n-1}. & \text{(3b)} \end{cases}$$

From that, we derive the recurrence equations, for $T_n = U_n/U_{n-1}$ and $Z_{i,n}$:

$$\begin{cases} T_n = \cosh\theta_n + \dfrac{Z_{i,n-1} \cdot \sinh\text{-}\theta_n}{Z_n} & \text{(4a)} \\[2ex] Z_{i,n} = \dfrac{Z_n}{\sinh\theta_n} (\cosh\theta_n - 1/T_n). & \text{(4b)} \end{cases}$$

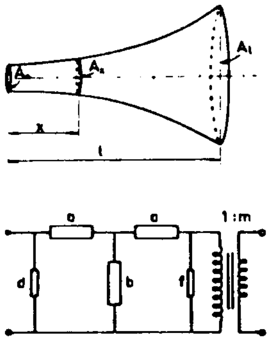
Starting from the radiation load, the recurrence equations involve (1) identify or calculate the load $Z_{i,n-1}$, (2) calculate U_n/U_{n-1}, (3) calculate $Z_{i,n}$, and (4) update the transfer U_n/U_0. This procedure is discussed in more detail in FANT (1985). One can verify that it is possible, when using these equations, to insert at any point in the cascade any quadripole of another type, for instance, parallel or series impedances (see Fig. 4), or a horn network.

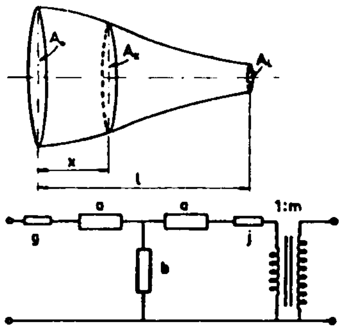
This is this latter method which has been implemented and used in our model.

1.3 DETERMINATION OF POLES AND ZEROS

Once we have obtained the transfer function of the vocal tract at any given frequency, the problem is to decompose this function into a quotient of two polynomials in the form of:

$$T(s) = \frac{\prod_{n=1}^{\infty} (s-s_n)(s-s_n^*) / (s_n \cdot s_n^*)}{\prod_{n=1}^{\infty} (s-\bar{s}_n)(s-\bar{s}_n^*) / (\bar{s}_n \cdot \bar{s}_n^*)} = \frac{A(s)}{N(s)}. \tag{5}$$





$$A(x) = \frac{A_0}{\cosh^2 \varepsilon} \cosh^2 \left(\frac{x}{h} + \varepsilon\right), \qquad A(x) = \frac{A_0 \cosh^2 \varepsilon}{\cosh^2 \left(\frac{x}{h} + \varepsilon\right)}$$

When $\varepsilon = 0$ the horn is catenoidal; when $\varepsilon = \infty$ the horn is exponential; and when $\varepsilon = \frac{x_0}{h} - j\frac{\pi}{2}$ with h approaching infinity, the horn is conical.

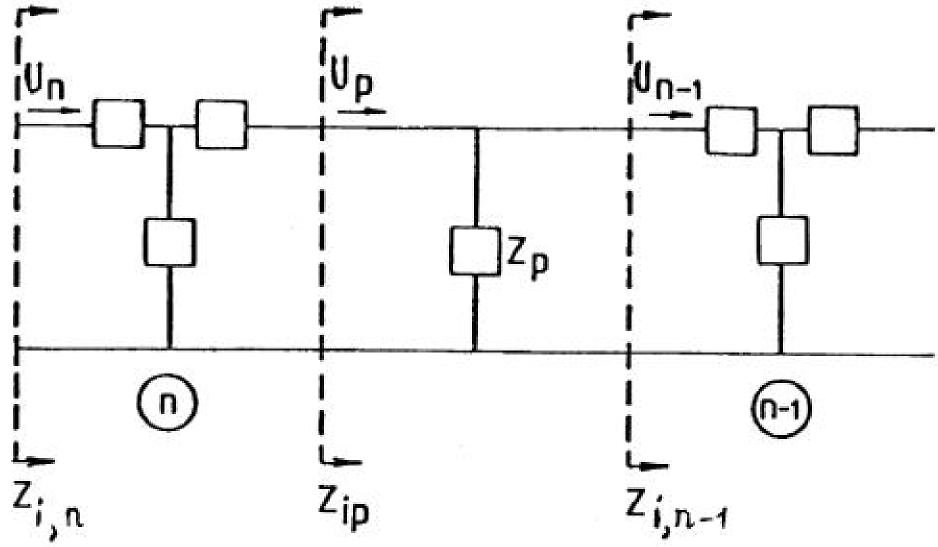
Resonator length	l
Nominal characteristic impedance	$Z_0 = \varrho c/A_0$
Nominal propagation constant	$\gamma_0 = \alpha + j\omega/c$
Horn cutoff frequency	$\omega_0 = c/h$
Propagation constant	$\gamma = \gamma_0(1-\omega_0^2/\omega^2)^{\frac{1}{2}} = \gamma_0 \cdot \tau$
Transfer constant	$\Gamma = l \cdot \gamma$
Characteristic impedance	$\begin{cases} Z = Z_0/\tau \text{ (increasing area)} \\ Z = Z_0 \cdot \tau \text{ (decreasing area)} \end{cases}$
Series element of T-network	$a = Z \,\mathrm{tgh} \frac{\gamma l}{2}$
Parallel element	$b = Z/\sinh \gamma l$
Transformer impedance ratio	$m^2 = A_0/A_l$

$$d = Z_0 h \gamma_0 \coth \frac{l}{h}, \qquad g = \frac{Z_0}{h\gamma_0} \coth \frac{l}{h}, \text{ and}$$

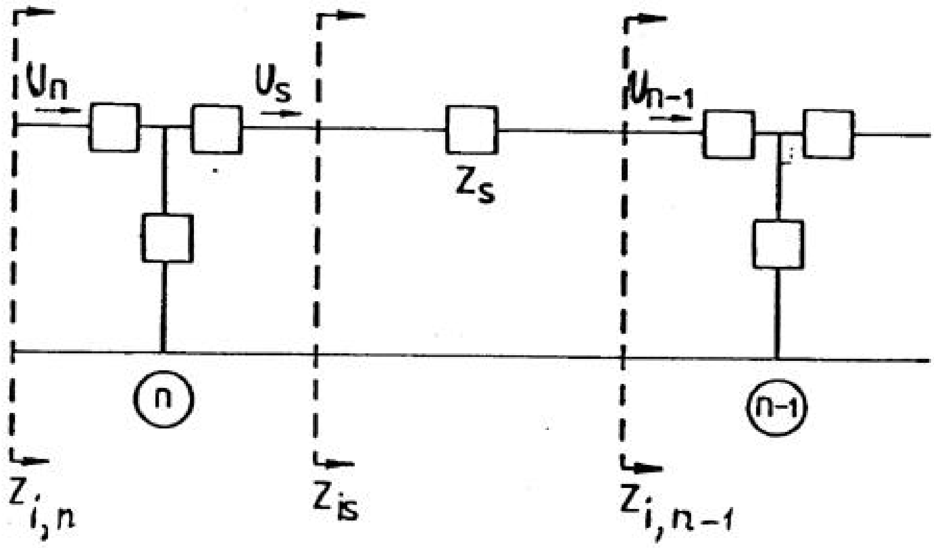
$$f = -Z_0 h \gamma_0 \coth\left(\frac{l}{h} + \varepsilon\right), \qquad j = -\frac{Z_0}{h\gamma_0} \,\mathrm{th}\left(\frac{l}{h} + \varepsilon\right).$$

	Increasing area		Decreasing area	
	d	f	g	j
Catenoidal horn	∞	$-Z_0 h \gamma_0 \coth \frac{l}{h}$	0	$\frac{-Z_0}{h\gamma_0} \,\mathrm{th} \frac{l}{h}$.
Exponential horn	$Z_0 h \gamma_0$	$-Z_0 h \gamma_0$	$Z_0/h\gamma_0$	$-Z_0/h\gamma_0$.
Conical horn	$Z_0 x_0 \gamma_0$	$-Z_0(x_0 + l)\gamma_0$	$Z_0/x_0\gamma_0$	$-Z_0/[(x_0 + l)\gamma_0]$

Fig. 3. Network representation of horn sections.



$$\begin{cases} Z_{i,p} = \dfrac{Z_p \; Z_{n-1}}{Z_p + Z_{n-1}} \\[2ex] T_p = \dfrac{U_{n-1}}{U_p} = \dfrac{Z_p}{Z_p + Z_{n-1}} \end{cases}$$



$$\begin{cases} Z_{is} = Z_s + Z_{in-1} \\ T_s = 1 \end{cases}$$

Fig. 4. Insertion of a parallel or of a series impedance.

In practice, we shall be looking only for a limited amount of poles and zeros, up to a given frequency limit.

1.3.1 Principle

We use one of the methods proposed by FANT (1960, p. 41) for dealing with all pole functions $T(s) = 1/N(s)$. The same method also applies to a zero function (these have to be dealt with separately, see Section 1.3.4). First we should notice that a function $T(s)$ containing poles only is, in the loss-less case, the inverse of a product of terms like $(1-s^2/s_n{}^2)$, with $s_n = j\omega_n$. This function has a real part only, and the poles correspond to the frequencies at which $1/T(s) = 0$. In the more general case, when the losses are not null but small, these pole frequencies are slightly modified. An imaginary part appears in $T(s)$ and we can then write:

$$N(s) = N_b(s) + jN_a(s) = 1/T(s). \tag{6}$$

The search for the poles involves two steps. The first is to determine the values of $s_n = j\omega_{n1}$ for which $N_b(s)$ equals zero. Then we use a first order approximation of $N(s)$ in the vicinity of its zero:

$$N(s) = N_b(s)_{s=j\omega_{n1}} + (s_n - j\omega_{n1})\cdot N'(s)_{s=j\omega_{n1}}. \tag{7}$$

$N(s_n) = 0$ leads then to

$$s_n = j\omega_{n1} - [N(s)/N'(s)]_{s=j\omega_{n1}}. \tag{8}$$

If we write

$$s_n = \sigma_n + j(\omega_{n1} + \Delta\omega_n), \tag{9}$$

and if we note that

$$N'(s)_{s=j\omega_{n1}} = \left[\frac{dN(s)}{ds}\right]_{s=j\omega_{n1}} = \frac{1}{j}\left[\frac{dN_b(s)}{d\omega} + j\,\frac{dN_a(s)}{d\omega}\right]_{s=j\omega_{n1}}, \tag{10}$$

the following results are obtained:

$$\begin{cases} \sigma_n = \dfrac{N_a \cdot N'_b}{N'^2_a + N'^2_b} & (11a) \\ \Delta\omega_n = -\sigma_n \cdot \dfrac{N'_a}{N'_b}\,, & (11b) \end{cases}$$

with

$$N_a = N_a(s)_{s=j\omega_{n1}}\,,\quad N'_a = \left[\frac{dN_a(s)}{d\omega}\right]_{s=j\omega_{n1}} \quad \text{and } N'_b = \left[\frac{dN_b(s)}{d\omega}\right]_{s=j\omega_{n1}}\,. \tag{12}$$

The principle of the program is thus to calculate $N_b(\omega)$ sampled with a certain frequency increment to look for changes of sign. Zero crossings are obtained by linear interpolation. The derivatives N_a' and N_b' are approximated by finite differences. In this way, we can determine σ_n and $\Delta\omega_n$.

1.3.2 Applications

In order to check the validity of the previous equations, we applied these equations to two simple cases: the transfer function of a uniform tube, which has the form :

$$T(s) = 1/\cosh(\alpha+s/c)l, \tag{13}$$

and a single second order resonance or low-pass filter, with

$$T(s) = \frac{s_2 \cdot s_2^*}{(s-s_2)(s-s_2^*)}\,. \tag{14}$$

The actual poles of the function, analytically solved from the zeros of $1/T(s)$ are:

$$s_1,\ s_1^* = \sigma_1 \pm j\omega_1 = -\alpha c \pm j\,\frac{\pi}{2}\frac{c}{l}\,, \tag{15}$$

for the uniform tube, and

$$s_2,\ s_2^* = \sigma_2 \pm j\omega_2 \tag{16}$$

for the single resonance.

The maximum amplitude frequencies of T(s) and the -3 dB bandwidths are for the uniform tube and the single resonance, respectively:

$$F_{m1} = \frac{1}{2\pi}\,,\quad B_{m1} \simeq -\frac{\sigma_1}{\pi} = \frac{\alpha c}{\pi}\,, \tag{17}$$

and

$$F_{m2} = \frac{1}{2\pi}\sqrt{\omega_2^2 - \sigma_2^2} \simeq \frac{\omega_2}{2\pi}\left(1 - \frac{1}{2}\left(\frac{\sigma_2}{\omega_2}\right)^2\right), \quad B_{m2} = -\frac{\sigma_2}{\pi}. \tag{18}$$

Finally, the method we want to test leads to the following results:

$$\begin{cases} F_1 = \frac{1}{2\pi}\left[\omega_1 + 0\right] & \text{(19a)} \\ B_1 = -\frac{1}{\pi}\left[-\frac{c}{l}\,\mathrm{th}\,\alpha l\right] - \frac{\alpha c}{\pi}, & \text{(19b)} \end{cases}$$

for the uniform tube, and

$$\begin{cases} F_2 = \frac{1}{2\pi}\left[\sqrt{\omega_2^2 - \sigma_2^2} - \omega_2\cdot\left(\frac{\sigma_2}{\omega_2}\right)^2\cdot\frac{1}{1+\left(\frac{\sigma_2}{\omega_2}\right)^2}\right] \simeq \frac{\omega_2}{2\pi}\left[1 - \frac{1}{2}\left(\frac{\sigma_2}{\omega_2}\right)^2\right] & \text{(20a)} \\ B_2 = -\frac{1}{\pi}\left[\sigma_2\cdot\frac{1}{1+\left(\frac{\sigma_2}{\omega_2}\right)^2}\right], & \text{(20b)} \end{cases}$$

for the single resonance. One can notice that, for the uniform tube, the three different estimates of the poles are identical. For the single conjugate pole resonance, the frequency of maximum amplitude differs from the pole frequency, and in fact, if we use the $\Delta\omega_n$ correction (which appear as second term in F1 and F2 expressions), our interpolation generates a pole frequency equal to the frequency of maximum amplitude. This correction is of the second order, and the discrepancy could be explained by the fact that, in order to derive this correction, we used only a first order expansion of T(s). A second order expansion would lead to very complex equations and has not been used.

In order to test this rather simple method, we have checked the results obtained with the Russian vowels (see definition in Section 3.1). Table 1 shows the comparison between the computed poles and bandwidths, and the values obtained by determining manually the frequencies corresponding to peak values of the absolute value of the transfer function, computed with a frequency increment of 1 Hz, and the -3 dB bandwidths. Moreover, we give the correction computed from Eq. 11. We can see that, in most of the cases, the results are better when this frequency correction is used. This a way to validate the method. An argument for the method is that an equal spacing of poles minimizes the

P1	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
N	535	534.9	0.0	0.1	0.0	32	32.0	0.0
ɑ	766	765.0	-0.1	10.7	-1.5	110	94.8	-13.8
e	460	460.0	0.0	0.4	0.0	36	35.5	-1.4
i	287	286.5	-0.2	0.6	-0.4	49	48.3	-1.4
o	604	603.1	-0.1	6.1	-1.1	76	68.3	-10.1
u	313	313.3	0.0	3.3	-1.0	52	52.0	0.0
ɨ	340	340.5	0.1	0.4	0.0	42	40.9	-2.6

P2	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
N	1522	1522.0	0.0	-1.4	0.0	63	63.7	+1.1
ɑ	1127	1129.8	0.2	-14.5	1.5	131	106.5	-18.7
e	1993	1992.8	0.0	8.9	-0.4	98	94.3	-3.8
i	2293	2293.3	0.0	8.9	0.3	85	81.8	-4.1
o	901	901.7	0.0	-6.5	0.7	67	59.4	-11.3
u	633	632.9	0.0	-2.5	0.4	37	36.9	-0.3
ɨ	1543	1542.8	0.0	0.1	0.0	79	78.3	-0.9

P3	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
N	2536	2536.5	0.0	-7.2	0.3	131	128.0	-2.3
ɑ	2473	2473.3	0.0	-2.1	0.0	85	83.9	-1.3
e	2834	2829.4	-0.2	17.7	-0.8	263	224.6	-14.6
i	3133	3128.6	-0.1	25.8	-0.9	261	204.3	-21.7
o	2394	2394.0	0.0	1.0	0.0	43	42.8	-0.5
u	2385	2385.5	0.0	1.0	0.0	26	26.3	+1.2
ɨ	2422	2422.1	0.0	1.6	0.0	61	60.2	-1.3

P4	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
N	3565	3567.2	0.0	-20.6	0.6	212	201.5	-5.0
ɑ	3629	3625.3	-0.1	15.6	-0.5	227	185.4	-18.3
e	3650	3650.0	0.0	6.3	-0.2	-	408.1	-
i	3723	3728.8	0.2	-15.8	0.6	298	237.3	-20.4
o	3494	3488.8	-0.1	26.7	-0.9	(250)	246.2	(-1.5)
u	3797	3750.8	-1.2	42.7	-2.3	(244)	172.3	(-29.4)
ɨ	3476	3475.5	0.0	1.6	0.0	103	99.4	-3.5

P5	(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)
N	4606	4611.9	0.0	-41.7	0.9	303	275.0	-9.2
ɑ	4090	4124.7	0.8	-66.9	2.4	(308)	259.0	(-15.9)
e	4194	4295.5	0.0	-104.8	2.5	(493)	396.2	(-19.4)
i	4786	4764.9	-0.4	9.7	-0.6	(608)	419.0	(-31.1)
o	4001	4005.4	0.1	-28.9	0.8	186	159.2	-14.4
u	4035	4041.4	0.2	-20.8	0.7	139	108.8	-21.7
ɨ	4199	4199.4	0.0	-6.3	0.2	166	156.5	-5.7

Table 1 : comparison, for the Russian vowels, between the pole frequencies and the bandwidths determined by the automatic and the manual methods.
(A) Pole frequencies, measured by manual pic detection (resolution : 1 Hz),
(B) Pole frequencies, automatically determined,
(C) Relative error of (B) in relation with (A) (in %),
(D) Correction applied to the N_b zero crossing frequency in order to get (B),
(E) Relative error in relation with (A), if no correction was used,
(F) -3dB bandwidth determined manually (a stroke - means that there are no points at -3dB from the pic, and the figures between parenthesis correspond to an approximation when there is only one of these points),
(G) Bandwidths, automatically determined,
(H) Relative error of (G) in relation with (F) (in %).

distance between pole frequency and maximum amplitude frequency. A shortcoming is when two poles merge to a single spectral peak. The bandwidth values do not check very well in some cases, but we cannot state that the -3 dB algorithm is the better one, since it might be affected by neighboring formants (leading to a non symmetrical peak or to a missing -3 dB point).

1.3.3 Loss-less case: interpolation method

In order to ensure a sufficient precision, the previous method must be used with a relatively small frequency increment, around 20 Hz, which means 250 points if we need the poles up to 5 kHz. This is not a great problem, especially if we need to plot the corresponding curve (see Fig. 16c, for instance) at the same time. But, if a rapid evaluation of the pole frequencies for a given vocal tract configuration is needed, another method can be used. We first calculate the values $y(f) = N_b(f)$ for increasing frequencies f_n with a relatively large frequency increment (a few hundreds Herz) until a sign reversal is found ($y_{n-1} \cdot y_n < 0$). We then compute the derivative

$$y'_n = \frac{y_n - y_{n-1}}{f_n - f_{n-1}} , \tag{21}$$

and the first estimate is

$$\begin{cases} F_1 = f_n + \Delta f_n = f_{n+1} & (22a) \\ \Delta f_n = -y_n / y'_n . & (22b) \end{cases}$$

The following estimates are also computed:

$$\begin{cases} f_{j+1} = f_j + \Delta f_j & (23a) \\ \Delta f_j = -y_j / y'_j , & (23b) \end{cases}$$

from $j = n$ to $j = n+k$, so that $\Delta f_{n+k} < 1$ Hz. The final pole frequency value $F_1 = f_{n+k+1}$ is now determined with an accuracy of about 0.1 Hz. In all, less than 40 points on the vocal tract transfer function is needed for the calculation of five formants (FANT, 1985).

1.3.4 Independent determination of poles and zeros

A useful procedure is to determine the pole part N(s) and the zero part A(s) of the transfer function

$$T(s) = \frac{A(s)}{N(s)} \tag{24}$$

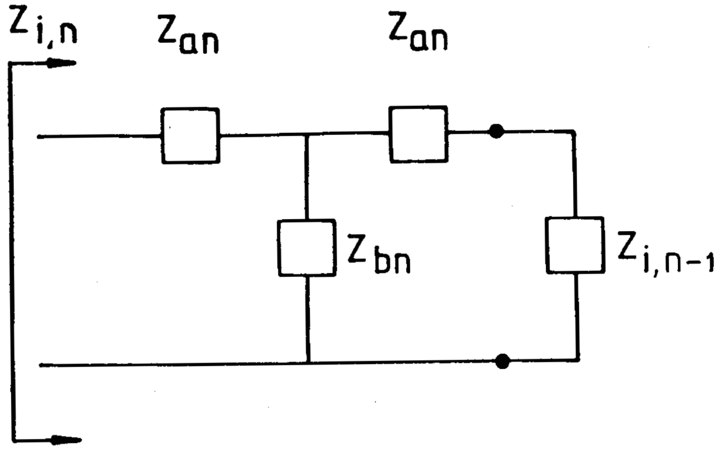


Fig. 5. Input impedance of a loaded quadripole.

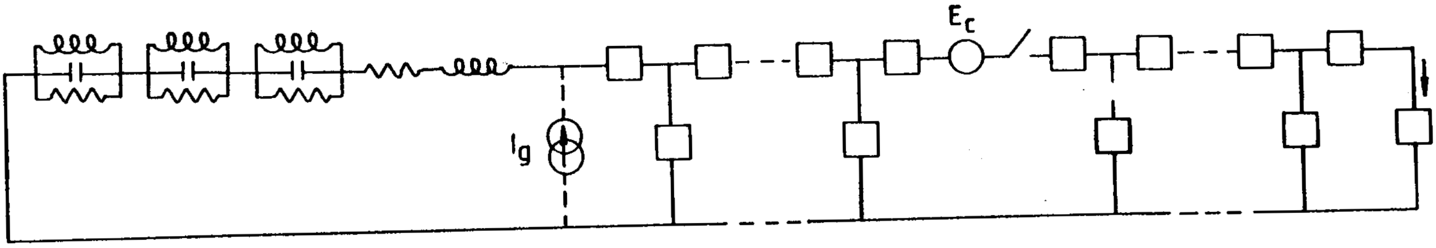


Fig. 6. Example of a consonantal vocal tract configuration.

separately from the analysis of the equivalent network. As illustrated in Fig. 6, we may introduce a glottal volume velocity source I_g providing an open circuit pressure E_c substituting the true consonant source E_c at the left of the switch in series with E_c. The transconductance I_0/E_c may now be broken up into two parts:

$$T = \frac{I_o}{E_c} = \frac{I_o}{I_g} \cdot \frac{I_g}{E_c} . \tag{25}$$

If there are no resonances causing short circuits in the shunting branches of the ladder network, both I_0/I_g and E_c/I_g will be all pole functions without zeros and may be calculated from the procedure in Section 1.3.1. The poles of E_c/I_g are the zeros of I_g/E_c and, thus, constitute the zero function A(s) of T(s). If shortcircuiting branches occur to the left of E_c only, there enter identical zero functions in I_0/I_g and E_c/I_g which evidently cancel in T(s). There remains in T(s) the pole function part of I_0/I_g and the pole function of E_c/I_g which is the zero function of T(s).

The procedure for evaluating the zero function A(s) is, thus, to start from E_c/I_g which is divided, i.e.,inverse filtered, by the product of shunting zero functions. The same procedure is applied to I_0/I_g for sorting out the pole function N(s). Shortcircuiting branches to the right of the consonantal source are included in this process of purifying I_0/I_g but do not affect the calculation of the zero function A(s).

This general approach is also useful for dealing with nasal consonants and nasalized vowels.

2. Vocal tract computation : losses and boundary conditions

In this part we analyze, one by one, the different losses in the vocal tract, and the boundary conditions. We describe different solutions (radiation and wall impedances), and try to assess their influence on the poles, zeros, and bandwidths of the vocal tract transfer functions.

2.1 RADIATION IMPEDANCE

In this section we first describe different models of radiation impedance, then present some comparison results.

2.1.1 Different models

The radiation impedance loads the end of the vocal tract line: this impedance contains a resistive part in which the radiated energy is consumed in series with a reactance representing the effective mass of vibrating air at the lips (FANT, 1960, p. 34).

2.1.1.1 Piston in sphere model (PIS)

The basic model commonly adopted is the model of a vibrating piston set in a spherical baffle; the equations have been derived by MORSE (1976, p. 323 ssq.); we use a computer subroutine written by U.K. LAINE (personnal communication). Fig. 7 shows the real and imaginary parts of the normalized radiation impedance $Z_R/(\rho c/A_0)$, where A_0 is the mouth area, for two different areas.

This model involves the calculation of series and is, thus, not very efficient computationally. For this reason we have looked into other models also.

2.1.1.2 STEVENS, KASOWSKI, & FANT model (SKF)

In 1953, STEVENS, KASOWSKI, & FANT reported the use of an analog electrical network to fit the theoretical model at low frequencies. This model is depicted in Fig. 8. It was incorporated in the LEA analog, FANT (1960). The normalized impedance is presented at Fig. 9 for two different mouth areas, in comparison with other models.

2.1.1.3 WAKITA & FANT model (WF)

Another model is the one used by WAKITA & FANT (1978). Their radiation impedance is defined by:

$$Z_R = R_0 + jX_0 \tag{26a}$$

with

$$R_0 = \frac{\rho\omega^2}{4\pi c} \cdot K_s(\omega) \tag{26b}$$

and

$$X_0 = \frac{\rho c}{A_0} \cdot \frac{\omega}{c} \cdot \frac{8}{3\pi}\sqrt{\frac{A_0}{\pi}} \tag{26c}$$

where $K_s(\omega)$ is defined by

$$\begin{cases} K_s(\omega) = 1 + \dfrac{0.6}{2\pi \cdot 1600} \cdot \omega \text{ , if } 0 \leqslant \ \leqslant 2\pi \cdot 1600 & (27a) \\ K_s(\omega) = 1.6 \text{ ,} \qquad \text{if } \omega \geqslant 2\pi \cdot 1600 \text{ .} & (27b) \end{cases}$$

For X_0, we use the approximation $\frac{8}{3\pi} = 0.8$, in order to conform to FANT's expression (1960, p. 36) of the "end correction" :

$$l_0 = 0.8\sqrt{\frac{A_0}{\pi}} \text{ .} \tag{28}$$

Fig. 9 shows the behavior of this model also.

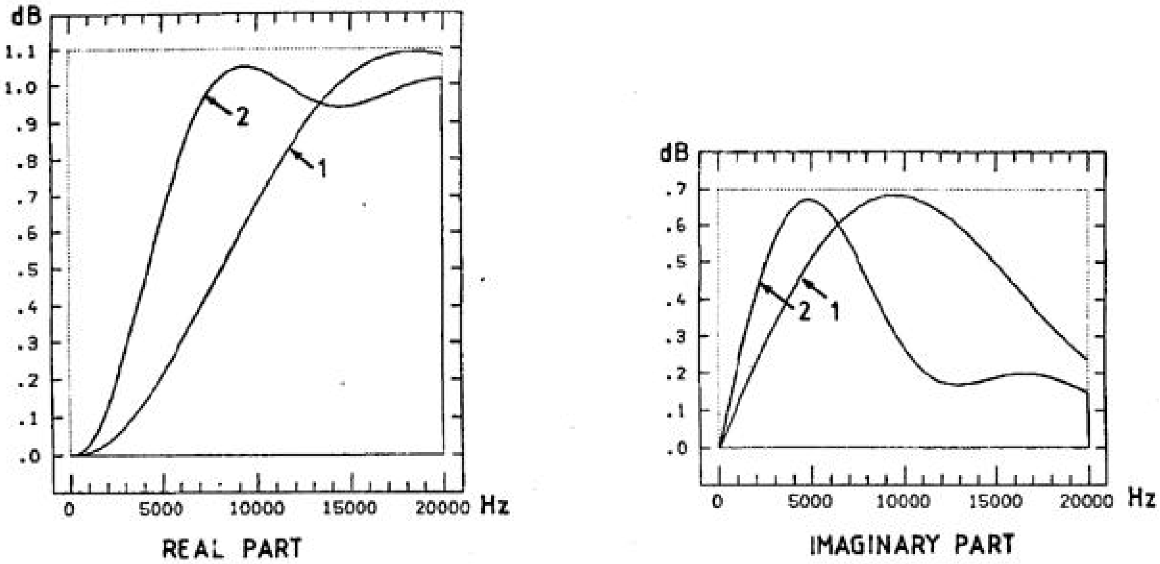
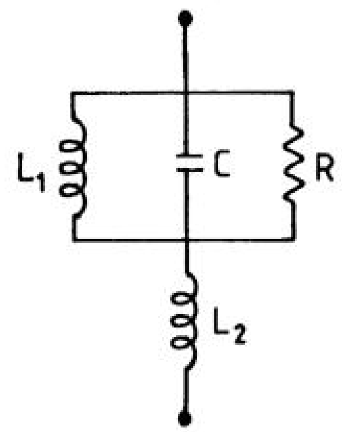


Fig. 7. Piston in Sphere normalized radiation impedance.
(1) Radiating area: 2 cm^2,
(2) Radiating area: 8 cm^2.



$L_1 = 4.23\ 10^{-4} \sqrt{A_o}$

$R = 45.9/A_o$

$C = 1.033\ 10^{-7} * A_o^{3/2}$

$L_2 = 7.11\ 10^{-5}/\sqrt{A_o}$

Fig. 8. Analog simulation of the radiation impedance.

2.1.2 Comparisons between the models

Fig. 9 shows that:

- for a relatively small area (2 cm^2), the SKF model fits the PIS model up to 10 kHz very well, whereas WF fits only up to 4-5 kHz ;
- for a large area (8 cm^2, which is not so common), the SKF model still fits PIS up to 4-5 kHz, when WF fits only up to 2-3 kHz.

Another way to compare the radiation models is to study their respective influences on the Russian vowels. Thus, we have computed the poles and the bandwidths for different conditions. We did not take neither the losses nor the wall impedances and the subglottal system into account. In the first comparison, we used only the imaginary part of the radiation impedance, and in the second one, the real part only. The results are given in the Tables IIa, IIb and IIc. Table IIa shows that the WF model provides pole frequencies with an error which can reach 2.3%, whereas the SKF model is much better, with a maximum error of 0.5%. For vowel bandwidths, the WF model is often better than the SKF model, but we can notice that the greatest relative error occurs for very small bandwidths, which means that their contribution to the total bandwidth is not important. Therefore, we can conclude that the SKF model is a very good approximation to the PIS model, with a much greater computational efficiency (around 7 times faster), and can be used in all practical cases.

As a remark, we should mention that the curve given by FANT (1960, p. 32) for the factor

$$K_s(\omega) = R_0 / \frac{\rho\omega^2}{4\pi c} \tag{29}$$

could be improved; we give the correct version in Fig. 10.

2.1.3 Influence of the radiation load on the poles and bandwidths

It is well known that the inductive part of the radiation impedance acts as to lengthen the vocal tract. Here, we study the separate influence of the resistive and inductive part of the radiation impedance on the poles and bandwidths of the Russian vowels.

From Table IIc, it is clear that the influence of the resistive part of the radiation impedance on the pole frequencies of the Russian vowels is very small and can be neglegted. On the other hand, the influence of the inductive part on the bandwidths is rather important, which may be explained with reference to L2 of Fig. 8.

2.1.4 Radiation transfer

The transfer from volume velocity at the lips to radiated pressure at a distance of a cm from the radiating surface at the lips is

		(A)	(B)	(C)	(B′)	(C′)
	F1	502.2	500.1	-0.4	501.4	-0.16
NT	F2	1508.4	1502.6	-0.4	1507.7	-0.05
	F3	2525.3	2511.2	-0.6	2523.5	-0.07
(6.0cm^2)	F4	3557.4	3527.8	-0.8	3552.5	-0.14
	F5	4605.0	4552.6	-1.1	4596.5	-0.18
	F1	642.3	640.2	-0.3	641.8	-0.08
[ɑ]	F2	1085.8	1081.6	-0.4	1085.4	-0.04
	F3	2470.4	2464.1	-0.3	2470.2	-0.00
(5.0cm^2)	F4	3622.2	3597.2	-0.7	3618.8	-0.09
	F5	4138.6	4132.3	-0.2	4137.6	-0.02
	F1	420.1	418.8	-0.3	419.5	-0.17
	F2	1973.1	1967.3	-0.3	1971.9	-0.06
[e]	F3	2815.0	2790.5	-0.9	2810.4	-0.16
(8.0cm^2)	F4	3645.8	3563.5	-2.3	3630.8	-0.41
	F5	4334.2	4246.0	-2.0	4318.7	-0.36
	F1	227.0	226.8	-0.0	226.9	-0.04
	F2	2275.8	2275.1	-0.0	2275.9	0.00
[i]	F3	3096.1	3070.6	-0.8	3095.8	0.00
(4.0cm^2)	F4	3731.6	3721.8	-0.3	3730.8	-0.02
	F5	4732.8	4649.4	-1.8	4721.9	-0.23
	F1	505.2	503.8	-0.3	505.1	-0.02
	F2	868.2	865.7	-0.3	868.6	+0.05
[o]	F3	2390.4	2388.6	-0.0	2391.0	+0.02
(3.2cm^2)	F4	3457.7	3456.7	-0.0	3457.8	0.00
	F5	4021.0	4017.7	-0.0	4020.9	0.00
	F1	237.4	237.4	0.0	237.8	+0.17
	F2	600.2	600.2	0.0	600.4	+0.03
[µ]	F3	2383.0	2383.1	0.0	2383.1	0.00
(0.65cm^2)	F4	3710.2	3710.2	0.0	3710.3	0.00
	F5	4055.9	4055.9	0.0	4056.1	0.00
	F1	289.6	289.1	-0.2	289.4	-0.07
	F2	1530.9	1518.2	-0.8	1529.0	-0.12
[ɨ]	F3	2414.1	2412.2	-0.0	2413.8	-0.01
(6.5cm^2)	F4	3472.1	3465.0	-0.2	3470.9	+0.03
	F5	4200.3	4192.4	-0.2	4199.1	-0.03

Table IIa. Effects of different models of radiation impedance on the pole frequencies of the Russian vowels.

(A) pole frequencies with the PIS model (imaginary part only),
(B) " " with the WF model (" "),
(C) difference (B)-(A)/(A), in %,
(B′) pole frequencies with the SKF model (imaginary part only),
(C′) difference (B′)-(A)/(A), in %.

Remarks :

- neither the losses nor the wall impedances and the subglottal system are taken into account ;
- the figures between parenthesis represent the radiation areas.

		(A)	(B)	(C)	(B′)	(C′)
	B1	3.2	3.5	+9.4	4.4	+37.5
NT	B2	38.6	42.1	+9.1	39.9	+3.4
	B3	112.7	116.7	+3.5	110.5	-1.9
(6.0cm^2)	B4	214.8	228.2	+6.2	213.8	+0.5
	B5	332.1	377.5	+13.7	343.5	+3.4
	B1	3.8	4.0	+5.3	4.9	+28.9
	B2	30.6	32.0	+4.6	33.6	+9.8
[a]	B3	61.1	62.5	+2.3	59.3	-2.9
(5.0cm^2)	B4	178.8	185.4	+3.6	174.8	-2.2
	B5	72.0	74.7	+3.7	71.3	-1.0
	B1	1.4	1.5	+7.1	2.0	+42.8
	B2	22.6	24.2	+7.1	22.9	+1.3
[e]	B3	67.8	72.6	+7.1	68.5	+1.0
(8.0cm^2)	B4	163.0	180.3	+10.6	168.6	+3.4
	B5	466.7	568.1	+21.7	502.5	+7.7
	B1	0.1	0.1	0.0	0.1	0.0
	B2	3.3	3.3	0.0	3.2	-3.0
[i]	B3	100.4	100.7	+0.3	95.7	-4.7
(4.0cm^2)	B4	57.1	57.8	+1.2	54.5	-4.5
	B5	126.5	131.7	+4.1	125.1	-1.1
	B1	2.0	2.2	+10.0	2.7	+35.0
	B2	18.5	19.1	+3.2	21.3	+15.1
[o]	B3	28.3	28.4	+0.4	27.0	-4.6
(3.2cm^2)	B4	11.0	10.8	-1.8	10.2	-7.2
	B5	28.2	28.0	-0.7	26.5	-6.0
	B1	0.2	0.2	0.0	0.2	0.0
	B2	0.3	0.3	0.0	0.4	+33.3
[u]	B3	0.3	0.3	0.0	0.2	-33.3
(0.65cm^2)	B4	0.4	0.4	0.0	0.3	-25.0
	B5	1.2	1.1	-8.3	1.1	-8.3
	B1	0.3	0.3	0.0	0.4	-33.3
	B2	102.2	110.3	+8.0	104.7	+2.4
[ɨ]	B3	21.5	22.4	+4.2	21.3	-0.9
(6.5cm^2)	B4	46.1	49.1	+6.5	46.0	-0.2
	B5	39.2	43.5	+11.0	40.1	+2.3

Table IIb. Effects of different models of radiation impedance on the bandwidths of the Russian vowels.

(A) bandwidths with the PIS model (real part only),
(B) " with the WF model (" "),
(C) difference (B)-(A)/(A), in %,
(B′) bandwidths with the WF model (" "),
(C′) difference (B′)-(A)/(A), in %.

Same remarks as for Table IIa.

		(A)	(B)	(C)	(D)	(E)	(F)
	F1	533.4	502.2	502.2	0.0	3.2	2.6
NT	F2	1600.2	1508.4	1507.9	0.0	42.1	26.9
	F3	2667.1	2525.3	2521.9	-0.2	116.7	80.5
($6.0cm^2$)	F4	3733.9	3557.4	3545.9	-0.5	228.2	141.1
	F5	4800.7	4605.0	4580.3	-0.8	377.5	204.0
	F1	676.1	642.3	642.3	0.0	3.8	3.9
[ɑ]	F2	1187.1	1085.8	1085.3	0.0	30.6	16.7
	F3	2554.0	2470.4	2469.1	0.0	61.1	40.4
($5.0cm^2$)	F4	3791.5	3622.2	3620.0	0.0	178.8	135.0
	F5	4185.9	4138.6	4134.5	-0.1	72.0	29.2
	F1	435.0	420.1	420.1	0.0	1.4	1.2
[e]	F2	2016.9	1973.1	1973.6	0.0	22.6	28.6
	F3	2912.7	2815.0	2819.4	+0.2	67.8	107.6
($8.0cm^2$)	F4	3856.8	3645.8	3650.7	-0.1	163.0	327.8
	F5	4632.0	4334.2	4253.0	-1.8	466.7	278.2
	F1	230.1	227.0	227.0	0.0	0.1	0.1
[i]	F2	2284.5	2275.8	2276.0	0.0	3.3	5.2
	F3	3290.1	3096.1	3104.6	+0.3	100.4	162.0
($4.0cm^2$)	F4	3800.7	3731.6	3729.5	0.0	57.1	56.5
	F5	4970.6	4732.8	4748.6	+0.3	126.5	384.9
	F1	535.1	505.2	505.2	0.0	2.0	1.9
[o]	F2	958.4	868.2	868.0	0.0	18.5	8.5
	F3	2437.1	2390.0	3290.0	0.0	28.3	15.6
($3.2cm^2$)	F4	3470.9	3457.7	3457.5	0.0	11.0	7.3
	F5	4050.0	4021.0	4020.2	0.0	28.2	19.9
	F1	248.8	237.4	237.3	0.0	0.2	0.1
[u]	F2	607.9	600.2	600.2	0.0	0.3	0.2
	F3	2384.3	2383.0	2383.0	0.0	0.3	0.2
($0.65cm^2$)	F4	3711.3	3710.2	3710.2	0.0	0.4	0.4
	F5	4059.2	4055.9	4055.9	0.0	1.2	1.1
	F1	295.0	289.6	289.6	0.0	0.3	0.3
[ɨ]	F2	1729.1	1530.9	1529.6	0.0	110.3	64.3
	F3	2436.6	2414.1	2413.5	0.0	22.4	10.1
($6.5cm^2$)	F4	3509.8	3472.1	3470.8	0.0	49.1	32.5
	F5	4227.1	4200.3	4199.1	0.0	43.5	30.7

Table IIc. Influence of the resistive and inductive part of the radiation impedance (PIS model) on the poles and bandwidths, for the Russian vowels.

(A) pole frequencies with a short circuit at the mouth,
(B) " " for $Re(Z_R) = 0$ and $Im(Z_R) \neq 0$,
(C) " " for $Re(Z_R) \neq 0$ and $Im(Z_R) \neq 0$,
(D) difference (B)-(A)/(A), in %,
(E) bandwidths for $Im(Z_R) = 0$ and $Re(Z_R) \neq 0$,
(F) " for $Im(Z_R) \neq 0$ and $Re(Z_R) \neq 0$.

Same remarks as for Table IIa.

essentially a differentiation. A more exact formula is given by FANT (1960):

$$P_a/V_o = \frac{\rho\omega}{4\pi a} K_T(\omega) \; . \tag{30}$$

The factor $K_T(\omega)$ is a smooth high frequency emphasis of about +1.5 dB per octave from 312 Hz to 5000 Hz. It represents the combined effects of directivity (baffle effect) and increase of radiation resistance in excess of the frequency proportionality. Because of lacking experimental verification $K_T(\omega)$ is generally omitted from calculations. However, for more detailed studies of the fricative generation, it should be taken into account together with the assumed initial frequency weighting of a random noise source.

2.2 VISCOSITY AND HEAT CONDUCTION LOSSES - COMPLEX CHARACTERISTICS

2.2.1 The formulas

The viscosity and heat conduction losses within a section of a cylindrical tube may, according to FANT (1960) and FLANAGAN (1972), be represented by a resistance R (viscosity) and a conductance G (heat conduction), see Fig. 2,

$$R = \frac{S}{A^2}\sqrt{\frac{\omega\rho\mu}{2}} \tag{31a}$$

and

$$G = S \cdot \frac{\eta - 1}{\rho c^2}\sqrt{\frac{\lambda\omega}{2C_p\rho}} \; , \tag{31b}$$

where η is the adiabatic gas constant, and μ, λ and C_p are the viscosity, the coefficient of heat conduction, and the specific heat of air.

From this and from Fig. 2, we can derive:

$$\alpha_R = \frac{1}{2A/S} \; \frac{\sqrt{\pi}}{c} \sqrt{\frac{\mu}{\rho}} \; \sqrt{f} \; , \tag{32a}$$

$$\alpha_G = \frac{1}{2A/S} \; \frac{\sqrt{\pi}}{c} \; (\eta - 1)\sqrt{\frac{\lambda}{C_p\rho}} \; \sqrt{f} \; , \tag{32b}$$

and

$$\alpha = \frac{1}{2A/S} \; \frac{\sqrt{\pi}}{c}\left[\sqrt{\frac{\mu}{\rho}} + (\eta - 1)\sqrt{\frac{\lambda}{C_p\rho}}\right]\sqrt{f} \; . \tag{32c}$$

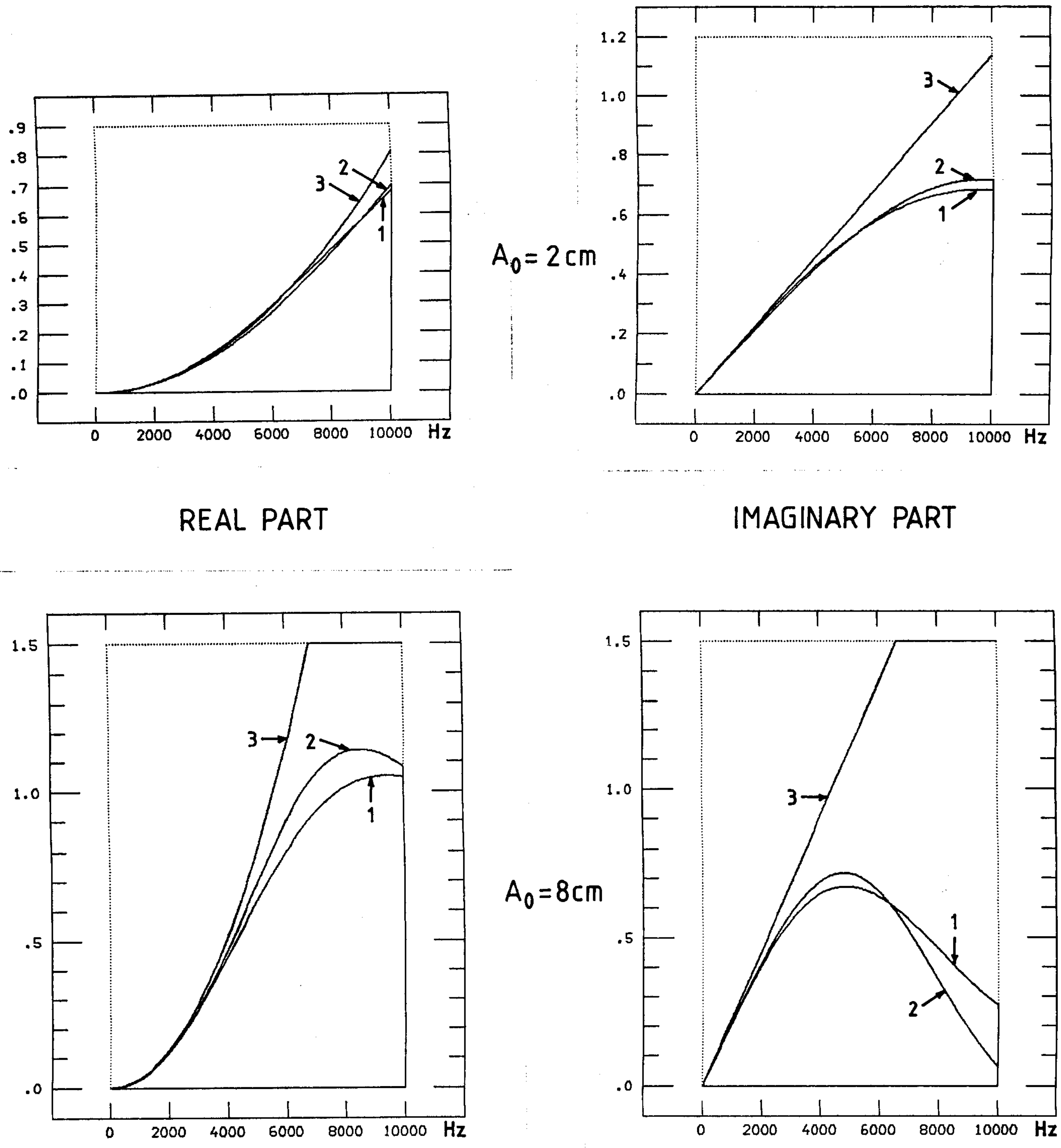


Fig. 9. Comparison of the radiation impedances for 3 models.
(1) Piston in Sphere model,
(2) Stevens-Kasowski-Fant model,
(3) Wakita-Fant model.

2.2.2 Complex characteristic impedance

In order to reduce the amount of computation, we have tested the effect on bandwidths of approximating the complex characteristic impedance

$$Z_c = Z_o \left[1 - j \cdot \frac{\alpha_R - \alpha_G}{\omega/c} \right] \tag{33}$$

by its loss-less real value Z_0, see Fig. 2. In Table III the results are shown. The differences are not very large and affect the low frequency poles only. Since at low frequencies the contribution from the viscous and thermal losses to the total bandwidth is rather small, one can conclude that a complex characteristic impedance is not a very important feature.

2.2.3 Influence of the viscous and thermal losses on the pole frequencies

We have also computed the pole frequencies retaining the viscous and thermal losses only and with Z_c real. These data differ from the loss-less case by less than 0.3 Hz over the entire range of the Russian vowels.

2.3 WALL IMPEDANCES

The finite vocal tract wall impedance as a factor affecting formant frequencies was first mentioned by van den BERG (1953) and by FANT (1960). Studies of diver's speech (FANT & SONESSON, 1964; FANT & LINDQVIST, 1968) contributed to the understanding of wall induced formant shifts. FLANAGAN (1972) derived equations for frequency and bandwidths changes but unfortunately he adopted data on human tissue which gave an order of magnitude too large bandwidth contributions compared to the more representative later data of ISHIZAKA et al. (1975) and of FANT et al. (1976). The correct order of magnitude had earlier been established by FANT (1972) by inference from the FUJIMURA & LINDQVIST (1971) resonance bandwidth data.

2.3.1 Distributed wall impedance models

The most common model (FLANAGAN, 1972; FANT, 1972; ISHIZAKA et al., 1975; MRAYATI, 1976; WAKITA & FANT, 1978), is to distribute the wall impedance uniformily over the tract. It was first shown by FANT (1972) that if the thickness of the walls is proportional to $A^{-0.5}$, the local resonance between the distributed capacitance $C(x)$ and the associated wall inductance $L_W(x)$ (see Fig. 11) is independent of $A(x)$ and equal to the closed tract resonance. If we consider the network in Fig. 11, we may write the complex expression of the associated propagation constant:

		(A)	(B)	(C)
	F1	4.8	4.8	0.0
	F2	8.3	8.3	0.0
NT	F3	10.7	10.7	0.0
	F4	12.6	12.6	0.0
	F5	14.3	14.3	0.0
	F1	10.4	10.5	1.0
	F2	11.3	11.9	5.3
[ɑ]	F3	19.0	19.0	0.0
	F4	20.6	20.7	0.5
	F5	26.1	27.2	4.2
	F1	5.6	5.9	5.3
	F2	11.6	11.2	-3.4
[e]	F3	16.0	15.8	-1.2
	F4	19.2	19.4	1.0
	F5	19.4	19.6	1.0
	F1	6.4	7.6	18.7
	F2	12.5	11.9	-4.8
[i]	F3	29.2	28.0	-4.1
	F4	21.7	21.8	0.5
	F5	22.0	21.8	-1.0
	F1	8.7	9.1	4.6
	F2	10.0	10.9	9.0
[o]	F3	15.0	14.8	-1.3
	F4	23.8	23.8	0.0
	F5	22.9	23.2	1.3
	F1	8.1	9.8	21.0
	F2	9.5	11.2	17.9
[u]	F3	19.3	18.9	-2.1
	F4	20.1	19.8	-1.5
	F5	21.3	21.2	-0.5
	F1	5.5	6.3	14.5
	F2	13.0	13.3	2.3
[ɨ]	F3	12.9	12.6	-2.3
	F4	19.7	19.3	-2.0
	F5	20.4	20.3	-0.5

Table III: Comparison between the bandwidths for Z_c real and Z_c complex (viscosity and heat conduction losses only, no radiation load).

(A) bandwidths for Z_c real,
(B) " for Z_c complex,
(C) difference (B)-(A)/(A), in %.

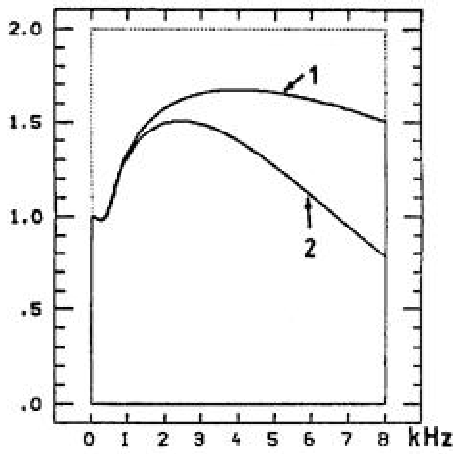


Fig. 10. Radiation resistance factor
(1) $A_o = 2\ cm^2$,
(2) $A_o = 8\ cm^2$.

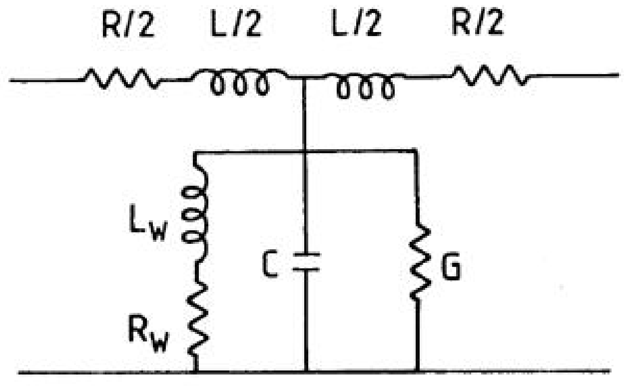


Fig. 11. T-cell with distributed wall impedance.

$$\gamma = \sqrt{(R + j\omega L)\ (j\omega C + G + \frac{1}{j\omega L_w + R_w})} \tag{34}$$

which, after some simplifications, leads to the approximation:

$$\gamma \simeq j\frac{\omega}{\omega_o}\sqrt{1 - \frac{F_w^2}{f^2}} - \left[\frac{1}{2}(\frac{R}{Z_o} + GZ_o) + \frac{\pi}{C_o}\ \frac{B_{wo}\cdot F_w^2}{B_{wo}^2+f^2}\right] \tag{35}$$

where

$$F_w = \frac{1}{2\pi\sqrt{L_w C}},\quad B_{wo} = \frac{R_w}{2\pi L_w},\quad C_o = \frac{1}{\sqrt{LC}},\ \text{and}\ Z_o = \sqrt{\frac{L}{C}}\,. \tag{36}$$

Thus, the wall impedance loading may be omitted and substituted by a frequency transformation such that the velocity of sound and the attenuation constant may be rewritten:

$$c = c_o(1 - F_w^2/f^2)^{-\frac{1}{2}} \tag{37a}$$

$$\alpha = \alpha_o + \frac{\pi}{C_o}\ \frac{B_{wo}\cdot F_w^2}{B_{wo}^2+f^2}\,. \tag{37b}$$

Since the formant frequencies are proportional to c, there follows that any formant frequency F_n can be derived from the infinite wall impedance case F_{ni} by:

$$F_n = (F_{ni}^2+F_w^2)^{\frac{1}{2}}\,. \tag{38a}$$

Furthermore, the bandwidth increase due to the walls is:

$$B_{nw} = B_{wo}(\frac{F_w}{F_n})^2\,. \tag{38b}$$

Even with the more common assumption of wall thickness, D being independant of A(x), Eq. 38a holds for the low frequency of the first formants.

It should be remarked that the network in Fig. 11 is not a complete representation of the wall impedance. The wall compliance and the wall radiation impedance have been omitted. We could use a radiation impedance if we want to mix in oral and nasal outputs as well as the sounds externally radiated by the walls of the tract, see FANT (1972).

2.3.2 Lumped wall impedance models

Our insight in the actual distribution of the wall impedance and its dependency of the cross-sectional area remains rudimentary. However, it is important to see to that the actual distribution adopted will provide closed tract resonances and bandwidths of the order of F_W = 180-200 Hz and B_W = 70-90 Hz, and about 20% higher values for an average female subject.

Since the wall effects are predominant in a low frequency region, it should be possible to attain the required overall effects by a few lumped impedance elements. This view was also supported by the observation of FANT et al. (1976), that the maximum vibrational amplitude measured externally was localized to one area just above the larynx and one at the cheeks. An implementation of this type is shown in Fig. 12. The model was tested by WAKITA & FANT who found that the assumption of a wall loading independence of the cross-sectional area A(x) gave rather too large F1 shifts for a vowel [ɑ]. We have, accordingly, modified the impedance level of both lumped branches by a factor $(6/A_m)^{0.5}$ where A_m is the mean area over a 4 cm part of the area function centered around the insertion points of the shunts. These two probe areas terminate at the lips and at the outlet of the larynx tube. The narrowing of the lower pharynx and larynx of [ɑ] is now matched by an increase of the wall impedance factor.

2.3.3 Calculations

We have made calculations of the formant frequency shifts in each of the six Russian vowels and a neutral single tube resonator when wall impedances are included by their imaginary terms only. Because of the high Q conditions, F_W/B_W = of the order of 2-3, this is allowable as found out from controls retaining the real part of the wall impedance which did not significantly influence the results (less than 1%).

Three different conditions were tested:

(1) lumped element representation according to Fig. 14 with area independant impedance level;

(2) the same with impedance level adjusted by the factor $(6/A_m)^{0.5}$;

(3) frequency transformation technique assuming distributed impedance with wall thickness proportional to $A^{-0.5}$, F_W = 190 Hz and B_W = 76 Hz.

		(A)	(B)	(C)	(D)	(E)
	F1	507.2	548.1	548.1	539.7	541.6
	F2	1523.3	1533.5	1533.5	1534.5	1534.8
NT	F3	2550.6	2554.2	2554.2	2557.4	2557.7
	F4	3593.7	3596.8	3596.8	3598.5	3598.7
	F5	4652.5	4657.0	4657.0	4656.4	4656.4
	F1	642.3	754.0	705.8	668.5	669.8
	F2	1085.8	1133.0	1109.6	1101.2	1102.3
	F3	2470.4	2473.1	2473.0	2477.5	2477.7
[ɑ]	F4	3622.2	3627.2	3626.3	3627.1	3627.2
	F5	4138.6	4162.8	4150.9	4142.9	4143.0
	F1	420.1	460.6	465.8	459.8	461.1
	F2	1973.1	1984.7	1985.7	1981.9	1982.2
[e]	F3	2815.0	2820.6	2820.2	2821.2	2821.4
	F4	3645.8	3652.2	3651.4	3650.6	3650.7
	F5	4334.2	4342.4	4342.5	4338.3	4338.4
	F1	227.0	289.6	304.6	295.3	296.0
[i]	F2	2275.8	2284.0	2285.3	2283.7	2283.7
	F3	3096.1	3117.3	3107.5	3101.6	3101.9
	F4	3731.6	3736.8	3735.1	3736.4	3736.4
	F5	4732.8	4743.8	4740.0	4736.4	4736.6
	F1	505.2	598.3	575.0	537.7	539.7
	F2	868.1	904.4	894.3	887.2	888.6
[o]	F3	2390.4	2393.0	2393.2	2397.8	2397.9
	F4	3457.7	3459.0	3458.8	3463.0	3462.9
	F5	4021.0	4026.0	4024.8	4025.4	4025.5
	F1	237.3	313.8	318.1	301.7	304.0
	F2	600.2	634.2	636.9	628.8	629.6
[u]	F3	2383.0	2384.7	2384.8	2390.6	2390.6
	F4	3710.2	3710.9	3710.9	3715.0	3715.0
	F5	4055.9	4058.3	4058.2	4060.4	4060.3
	F1	289.6	342.2	359.0	345.5	346.4
[ɨ]	F2	1530.9	1541.5	1541.6	1541.4	1542.6
	F3	2414.1	2418.8	2420.3	2421.5	2421.6
	F4	3472.1	3473.3	3473.4	3477.2	3477.3
	F5	4200.3	4201.6	4201.8	4204.6	4204.6

Table IVa. Pole frequencies for the Russian vowels, with different yielding wall models.

(A) no walls,
(B) area independant lumped wall impedances (imaginary part only),
(C) area dependant lumped wall impedances (imaginary part only),
(D) frequency transformation $c = c_0.(1-F_W^2/f^2)^{-1/2}$, with F_W = 190 Hz,
(E) formula $F_n^2 = F_{ni}^2 + F_W^2$.

Remark : the radiation impedance is the imaginary part of the PIS model.

		(A)	(B)	(C)	(D)
	F1	10.5	10.5	8.2	9.2
	F2	1.0	1.0	1.1	1.2
NT	F3	.2	.2	.4	.4
	F4	.1	.1	.2	.2
	F5	.1	.1	.1	.1
	F1	17.4	12.3	5.5	6.1
	F2	7.2	3.4	2.0	2.3
[ɑ]	F3	.2	.2	.4	.4
	F4	.2	.2	.2	.2
	F5	.9	.5	.2	.2
	F1	12.3	13.5	11.2	12.6
	F2	.9	1.0	.7	.7
[e]	F3	.3	.3	.3	.3
	F4	.3	.2	.2	.2
	F5	.3	.3	.1	.1
	F1	27.9	31.8	22.4	29.5
	F2	.5	.6	.5	.5
[i]	F3	1.0	.5	.3	.3
	F4	.2	.1	.2	.2
	F5	.4	.2	.1	.1
	F1	18.4	15.6	8.0	9.3
	F2	6.5	4.8	3.1	3.5
[o]	F3	.2	.2	.5	.5
	F4	.1	.0	.2	.2
	F5	.2	.1	.2	.2
	F1	29.8	30.4	21.4	28.3
	F2	8.5	9.4	6.4	6.9
[u]	F3	.1	.1	.5	.5
	F4	.0	.0	.2	.2
	F5	.1	.1	.2	.2
	F1	21.1	25.9	17.9	21.9
	F2	1.0	1.0	1.0	1.2
[ɨ]	F3	.3	.4	.5	.5
	F4	.1	.1	.2	.2
	F5	.0	.1	.2	.2

Table IVb. Bandwidths for the Russain vowels with different yielding wall models.

(A) area independant lumped wall impedances (complex values),
(B) area dependant lumped wall impedances (complex values),

(C) frequency transformations $c = c_0.(1-F_W^2/f^2)^{-1/2}$,

and $\alpha = \alpha_0 + \frac{\pi}{c_0} \cdot \frac{B_{WO} \cdot F_W^2}{B_{WO}^2 + f^2}$

with F_W = 190 Hz and B_{WO} = 76 Hz.

(D) formulas $F_n^2 = F_{ni}^2 + F_W^2$ and $B_{nW} = B_{WO} \cdot (\frac{F_W}{F_n})^2$.

Remark : the radiation impedance is the imaginary part of the PIS model.

As seen from Table IVa, model (3) provides the overall smallest shifts and model (1) the largest shifts. Alternative (2) is recommended as a representative and simple model to implement.

The bandwidth increase associated with the resistive part of the wall impedance is documented in Table IVb. The effect is greatest for B1 of vowel [i] and [u] (30 Hz). The area independence leads to an apparent increase in the [o] and [ɑ] bandwidths.

The simple formula Eq. 38a and 38b of transformed frequency and bandwidth, case (3), apparently have a good predictive potential.

We may conclude that there is an apparent need for new and more detailed experimental data on vocal tract wall impedance and its distribution.

2.4 GLOTTIS IMPEDANCE

2.4.1 The equations

Since the model we deal with is a frequency domain model, it is not possible to attain a complete simulation of the time-varying glottal impedance; thus, we define a frozen state for the glottis and we study the frequency properties of the vocal tract for this given state. The glottis is modeled as a rectangular slit of the area A_g and the length l_g; the depth or thickness is d. In order to complete the model, we must add the pressure drop P_s at the glottis. According to van den BERG et al. (1957) and FANT (1960, p. 267), the glottis flow resistance R_F can be decomposed into two terms $R_F = R_L + R_T$, with

$$R_L = \frac{12\mu d}{A_g^3/l_g^2}, \tag{39}$$

corresponding to a laminar streaming resistance, and R_T being the term due to turbulent losses, associated with the BERNOUILLI equation:

$$\begin{cases} P_s = k_g \cdot \dfrac{\rho v_g^2}{2} & \text{(40a)} \\[2ex] R_T = \dfrac{P_s}{u_g} = k_g \dfrac{\rho v_g}{2A_g}, & \text{(40b)} \end{cases}$$

where $v_g = u_g/A_g$ is the particle velocity of the glottal flow.

From the differentiation of Eq. 40a we see that the differential resistance, appropriate for the small signal analysis, is $R_D = 2R_B$.

The glottal inductance is:

$$L_g = \frac{\rho d}{A_g}. \tag{41}$$

Thus, the glottis impedance is:

$$Z_g = \left(\frac{12\mu d}{A_g^3/l_g^2} + \frac{\sqrt{2k_g \rho P_s}}{A_g}\right) + j\frac{d}{A_g} . \tag{42}$$

van den BERG et al. (1957) measured $k_g = 0.785$, and ANANTHAPADMANABHA & FANT (1982) proposed $k_g = 1.1$.

Eq. 42 is implemented in our model, and the factor k_g can be given any value.

2.4.2 Influence of the glottis impedance on the vocal tract

Table V gives the pole frequencies for a uniform tube for different types of glottis impedance. The subglottal system is removed and replaced by a short-circuit; there are no vocal tract losses, and the radiation impedance is short-circuited.

We can check that the pole frequencies are independent of the glottis resistance, if there is no glottis inductance. On the other hand, the shift caused by the inductance is insignificant if the glottis resistance is used. Finally, the bandwidths are very much affected by the inductance.

FANT & LILJENCRANTS (1979) have shown that the main correlate of perceived formant intensity is the mean value of the corresponding formant oscillation during a fundamental period. If the glottal closed state occupies more than one half of the period, the mean value of the formant amplitude will not change much in comparison with a moderate and a very large damping in the open phase. With a very large damping in the open phase, it is the relative duration of the closed phase rather than the degree of damping which determines the formant amplitude and, thus, indirectly a measure of the effective bandwidth. Therefore, the frozen state calculations of the glottal contribution to formant bandwidths are not a reliable substitute for an exact interactive calculation procedure, see further ANANTHAPADMANABHA & FANT (1982) and FANT & ANANTHAPADMANABHA (1982).

For the present study we have selected $P_s = 16$ cmH_2O, $A_g = 0.027$ mm^2, $l_g = 1.2$ cm, $d = 0.3$ cm, and $k_g = 1$, which represent a rather small glottal opening. The calculated $R_F = 49\Omega$ and $R_D = 221\Omega$ in all acoustical ohms lead to a total resistance of $6.7\,\rho c$. The results of the bandwidth calculations are reported in Table VI. The effect of the glottal load is largest on formants that are associated with the back cavity.

2.5 THE SUBGLOTTAL SYSTEM

2.5.1 Description

When the glottis is open, its impedance is no longer very high, and thus, the coupling between the lungs and the vocal tract is no longer negligible, for instance, it happens that some extra resonances appear in the spectra of certain sounds (FANT, 1969; FANT et al., 1972). It

$Re(Z_g)$	$Im(Z_g)$	F1	F2	F3	F4	F5	B1	B2	B3	B4	B5
infinite		533.4	1600.2	2667.1	3733.9	4800.7	0.0	0.0	0.0	0.0	0.0
0	1	582.2	1617.9	2677.8	3741.5	4806.7	0.0	0.0	0.0	0.0	0.0
1	0	533.4	1600.2	2667.1	3733.9	4800.7	16.9	16.9	16.9	16.9	16.9
1	1	534.7	1603.5	2671.2	3738.1	4804.7	16.5	13.8	10.4	7.6	5.6

Table V. Effect of different types of glottis impedance on the neutral tube.

No losses and no inductances, except for Z_g.

$l = 17.54$ cm $\quad$ $R_g = 270.0\ \Omega$

$A = 6$ cm^2 $\quad$ $L_g = 12.68$ mH.

	(A)	(B)	(C)	(D)	(E)	(F)
	533.4	535.0	534.7	0.2	14.7	11.1
	1600.2	1603.1	1603.5	0.2	14.2	10.7
NT	2667.1	2671.1	2671.2	0.1	11.0	8.2
	3733.9	3738.1	3738.1	0.1	7.8	5.9
	4800.7	4804.7	4804.7	0.0	5.7	4.3
	676.1	682.8	686.7	1.0	79.3	73.2
	1187.1	1191.0	1190.6	0.3	28.8	27.1
[ɑ]	2554.0	2558.5	2558.6	0.2	11.9	12.6
	3791.5	3809.6	3809.5	0.0	26.7	27.5
	4185.9	4224.1	4224.1	0.9	68.9	70.7
	435.0	436.2	435.8	0.3	13.1	12.4
	2016.9	2027.0	2027.8	0.5	31.8	29.0
[e]	2912.7	2936.7	2936.9	0.8	52.2	54.5
	3856.8	3885.6	3885.5	0.7	50.4	51.8
	4632.0	4645.5	4645.5	0.3	21.1	21.6
	230.1	230.5	230.4	0.2	11.5	11.4
	2284.5	2295.6	2296.7	0.5	31.9	35.0
[i]	3290.1	3300.1	3300.1	0.3	18.0	18.7
	3800.7	3830.3	3830.3	0.8	55.4	57.1
	4970.6	4979.6	4979.6	0.2	12.8	13.1
	535.1	540.4	539.5	1.0	46.2	40.7
	958.4	959.8	959.9	0.1	12.1	11.8
[o]	2437.1	2439.3	2439.4	0.0	5.9	6.3
	3470.9	3517.2	3517.0	1.3	77.2	79.8
	4050.0	4073.7	4073.7	0.6	44.7	45.9
	248.8	249.5	2493.3	0.3	14.3	14.2
[u]	607.9	608.6	608.7	0.1	10.6	9.3
	2384.3	2385.4	2385.5	0.0	3.1	3.4
	3711.3	3750.2	3749.8	1.0	57.1	58.7
	4059.2	4078.7	4078.7	0.5	36.8	37.8
	295.0	295.7	295.4	0.2	11.9	11.8
	1729.1	1729.5	1729.5	0.0	1.3	1.3
[ɨ]	2436.6	2442.9	2443.3	0.3	17.3	18.5
	3509.8	3519.5	3519.5	0.3	17.3	17.9
	4227.0	4246.7	4246.6	0.5	30.9	31.7

Table VI. Influence of the subglottal system on the Russian vowels.

(A) pole frequencies in the lossless case,
(B) pole frequencies with glottis + subglottal system,
(C) pole frequencies with the glottis only,
(D) difference (B)-(A)/(A), in %,
(E) bandwidths with glottis + subglottal system,
(F) bandwidths with the glottis only

All losses and inductances are null, except for the subglottal system ; P_s = 16 cmH_2O, A_g = 0.027 cm^2 and l_g = 1.2 cm, that is R_g = 270 Ω and L_g = 12.67 mH.

is, thus, important to have a model of this subglottal system, and to connect it to the vocal tract electrical line.

FANT et al. (1972) have derived a model of the subglottal system and resonance properties of its input impedance. We represent this input impedance by three RLC modules as specified by ANATHAPADMANABHA & FANT (1982) (see Fig. 13). The frequency characteristics of this impedance is shown in Fig. 14. It is clear from this figure that there are three poles, and three zeros, of which the first one is at the zero frequency (the lung volume capacitor approximated by a shortcircuit).

2.5.2 Influence on the vocal tract

The influence of the subglottal system on the poles and zeros of the vocal tract is rather complex. It is obvious that this influence depends very much on the glottis impedance and also on the position of the pressure source.

In fact, since the glottal impedance in normal voicing is relatively high, the role of the subglottal system is small. It is possible to verify this assumption from Table VI, where we can see that the major modification to the poles and the bandwidths is due to the glottis impedance itself and not to the subglottal system.

2.6 CONSTRICTION

To compute the resistance at a constriction in the vocal tract, we first compute the DC glottal air flow from the lung pressure and the overall aerodynamic resistance, the main part of which resides either in the glottis or in the supraglottal constriction, compare Section 2.4.1. We assume that the volume velocity flow is the same in the whole vocal tract. The average particle velocity in a constriction of the cross-sectional area A_c, is thus $v_c = u_c/A_c$.

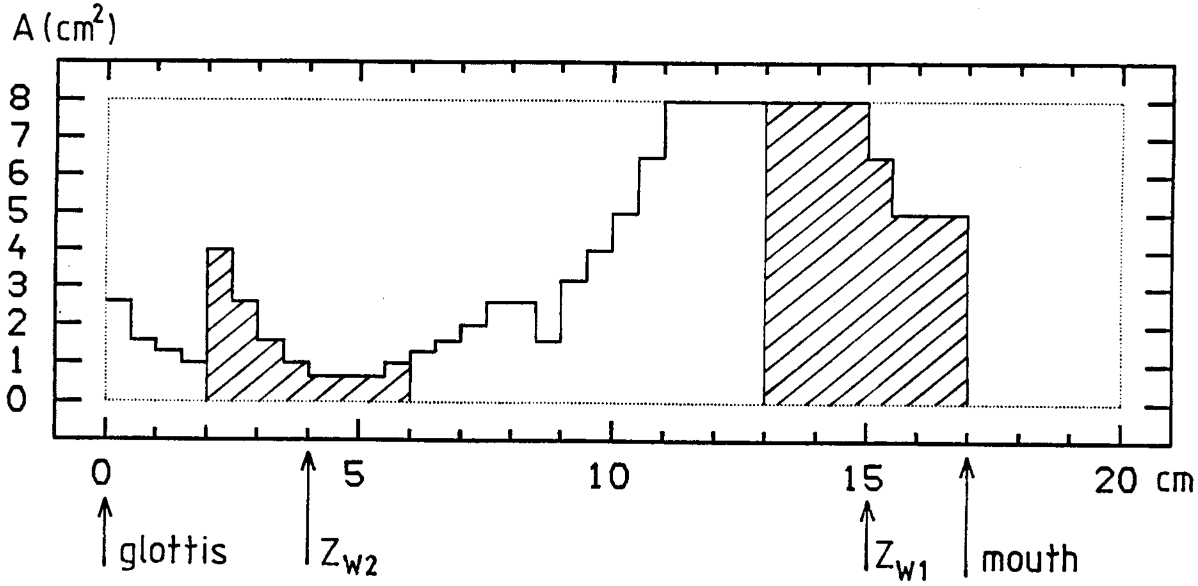
According to FANT (1960, p. 173-174), there are three different sources of resistance. The first one corresponds to the BERNOUILLI equation, written with the coefficient k_c (see Eq. 40a), for a small signal approximation:

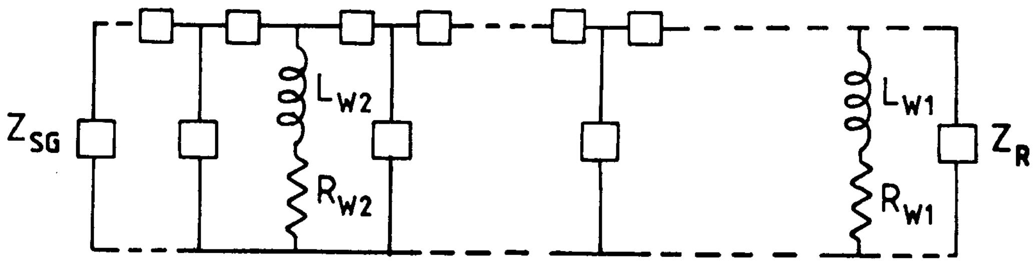
$$R_1 = 2(k_c \cdot \frac{\rho v_c}{2A_c}) \tag{43}$$

The second term comes from the laminar streaming:

$$R_2 = (\frac{\Delta P}{\Delta l})_L \cdot \frac{l_c}{u_c} = \frac{32\mu l_c}{A_c d_c^2} , \tag{44}$$

where l_c is the length and d_c the diameter of a constricted (cylindrical) tube.

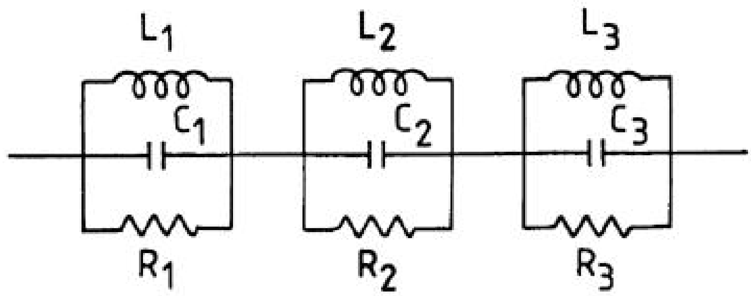




L_{w2} = 0.015 H
R_{w2} = 7.16 Ω

L_{w1} = 0.038 H
R_{w1} = 18 Ω

Fig. 12. Lumped impedances yielding walls representation.



	Subglottal formants	Bandwidths	Inductances
F1	615 Hz	246 Hz	3.80 mH
F2	1355 Hz	155 Hz	0.72 mH
F3	2110 Hz	140 Hz	0.27 mH

L1 = 3.80 mH	C1 = 17.6 μF	R1 = 36.7 Ω
L2 = 0.72 mH	C2 = 19.2 μF	R2 = 53.6 Ω
L3 = 0.27 mH	C3 = 21.1 μF	R3 = 53.9 Ω

Fig. 13. Subglottal impedance network.

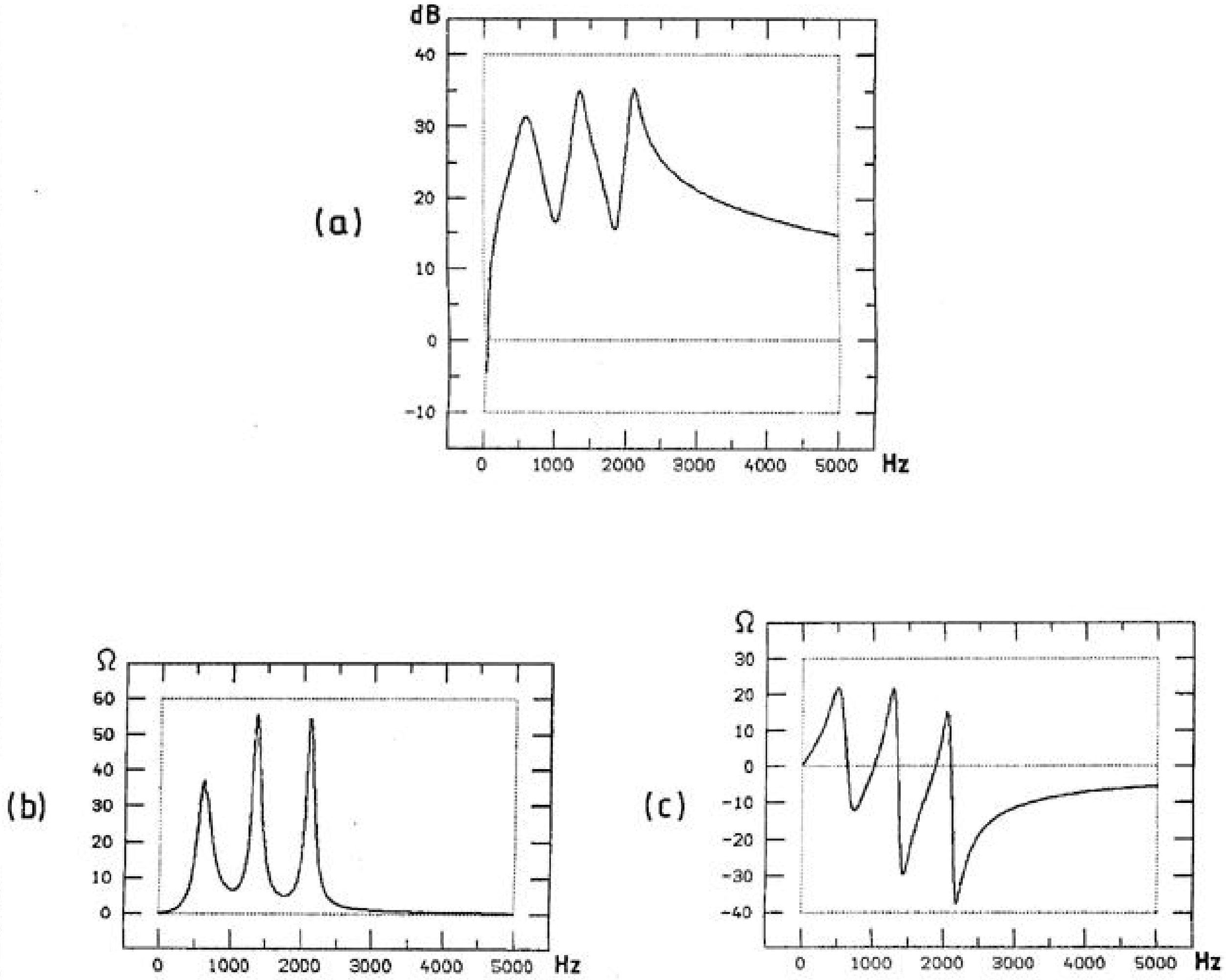


Fig. 14. Subglottal input impedance.
(a) amplitude,
(b) real part,
(c) imaginary part.

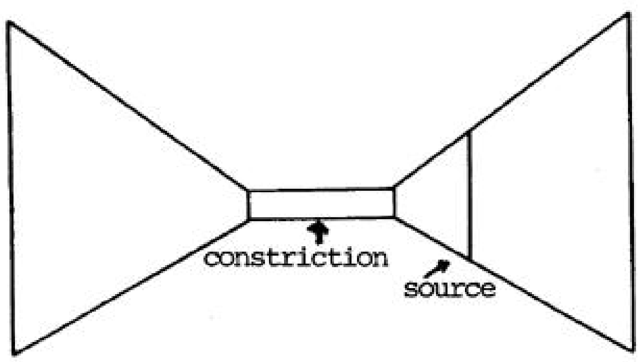


Fig. 15. Example of the representation of a construction by conical hans.

The third term, corresponding to the turbulent flow, is:

$$R_3 = \left(\frac{\Delta P}{\Delta l}\right)_T \cdot \frac{l_c}{u_c} = \frac{\xi \cdot \rho}{d_c \cdot 2} \cdot v_c^2 \cdot \frac{l_c}{v_c A_c} , \tag{45}$$

with

$$\xi = 0.3164\ R_e^{-1/4} , \tag{46}$$

and

$$R_e = \frac{v_c d_c}{\nu} \tag{47}$$

being the REYNOLDs number for the constriction; ν is the kinematic coefficient of viscosity, defined as the ratio of the viscosity coefficient to the density of the gas.

In fact, for constriction sizes usually encountered in speech production, the laminar streaming term is negligible compared to the one corresponding to the flow turbulence in a tube, and this latter is usually smaller than the BERNOUILLI term. More specifically, the nature of the resistance depends on the type of flow in the tube and, thus, a choice should be made between the laminar and the turbulent terms, depending on the REYNOLDs number at the constriction.

The Rl resistance is related to the compression of the flowing air veine from a large area to a small one: it can, thus, be located at the upstream inlet of the constriction. The turbulent term corresponds to a resistance which is spread along the constricted tube, but, since it is somewhat smaller than the BERNOUILLI resistance, it can be lumped into that. Thus, we will use only one resistance at the inlet of the constriction. We determine the tube wich has the smallest area, and insert this resistance in series between the T-cell corresponding to this constriction tube and the previous one.

2.7 HORN SECTIONS

We have already given the network representation for a rather general type of horn sections. Some preliminary experiments showed that it is important to define the vocal tract shape in the vicinity of a consonantal source with a great precision, if we want to be able to study the effects of a shift of the source location. If we want to avoid an exceedingly large amount of small tube sections to represent the tract, the horn sections seem to be a good solution to this problem. As a practical modular unit, we decided to use conical horns, and to define the constriction in the way shown by Fig. 15. Another alternative, which is simpler for overall cavity shape modeling, is the catenoidal horn. It is less suited for multiple cascading to study the influences of the source locations.

The equations for the conical horn are obtained by using

$$\varepsilon = \frac{x_o}{h} - j\,\frac{\pi}{2} \tag{48}$$

with h going to infinity, in the equations of Fig.3. We use only the increasing area model in both forward and backward directions depending on the end areas of the cone.

In order to introduce the viscosity and heat conduction losses, we have defined an average value of α.

We know that the area function is:

$$A(x) = A_o\left(\frac{x_o + x}{x_o}\right)^2 . \tag{49}$$

For a very small section, dx long, we have:

$$\alpha(x) = \frac{K\sqrt{f}}{\sqrt{A(x)}} . \tag{50}$$

Thus, we can define

$$\bar{\alpha} = \frac{1}{l}\int_0^l \frac{K\sqrt{f}}{\sqrt{A(x)}}\,dx = K\sqrt{f}\cdot\frac{x_o}{\rho A_o}\cdot \mathrm{Log}\left(\frac{x_o + l}{x_o}\right) . \tag{51}$$

This average value is used in the T-cell part of the network in Fig. 13. We have checked for several configurations that approximations with 0.1 cm long tube sections provide pole frequencies within 1 Hz of the conical horn models and bandwidths within 10-20%, depending on the cone angle (the fit is better for small angles). This accuracy is judged to be satisfactory.

3. Applied studies

In this part we report on vocal tract simulations applied to the reference Russian vowels. Then, we give the results of a preliminary study on one constrictive consonant, and, finally, provide some data on helium speech.

3.1 REFERENCE DATA FOR THE RUSSIAN VOWELS

By "Russian vowels" we mean the set of vocal tract configurations corresponding to the articulation of the Russian vowels [ɑ],[e],[i],[o], [u], and [ɨ] studied in FANT (1960). The first line (x = 0) of the table on p.115 must be read as the mouth radiating area for the vowels. The second line corresponds to the first 0.5 cm long section, and so on, going from the mouth towards the glottis. We should notice that the vocal tract for the vowel [ɨ] is only 18.5 cm long and is terminated at the glottis by 4 sections of the 3.2 cm^2 area.

To this set of vowels we have added a "neutral" tube or "uniform" tube, noted NT, defined as a 16.5444 cm long tube of constant area 6 cm^2. With a radiation inductance of $0.8\sqrt{A_0/\pi}$ and infinite wall impedance, the resonance frequencies become odd multiples of 500 Hz. In Table I, the configuration noted N corresponds to l = 16.37 cm and A = 8 cm^2.

3.1.1 Reference formant frequencies

Table VII lists the formant frequencies for the Russian vowels, assuming closed glottis state and both infinite wall impedance and area-dependent lumped wall impedances. These data check relatively well with the FANT (1960) and WAKITA & FANT (1978) data, considering that the radiation models are different, and that we apply a correction to the lumped wall impedances.

3.1.2 Incremental bandwidths

Table VIII gives the contribution to the total bandwidths of the different losses in the tract. These data are of the same order of magnitude as the FANT (1960) data. Line (4), corresponding to the sums of lines (1), (2), and (3), and line (5), corresponding to the bandwidths when using viscous and thermal, radiation and wall losses together, are in good agreement: the "additivity" property of the contributions to the overall bandwidths is verified.

The effect of glottis inductance on glottal dissipation is to reduce the glottal bandwidth component at higher frequencies, where the imaginary part of the glottal impedance is no longer negligible compared with the real part.

3.2 PRELIMINARY STUDY OF A CONSTRICTIVE

The purpose of this section is to illustrate simulations from a given vocal tract configuration and aerodynamics. We use the configuration [š], i.e., [ʃ] from FANT (1960, p. 172), see Fig. 16a.

3.2.1 The general conditions

We define here all the conditions which have been kept constant throughout the study. We have used viscous and thermal losses with a shape factor of 1. The radiation model was the Piston in Sphere with the +6 dB/oct radiation transfer. The wall impedances were simulated by the frequency transformations (Eqs. 37a and 37b). The glottis state was defined by l_g = 1.2 cm, d = 0.3 cm, P_s = 8 cmH_2O, and k_g = 1. No subglottal system was used. The constriction was modeled as a tube of 0.4 cm length and 0.4 cm^2 area, with a pressure source located at the outlet of the tube; the corresponding resistance was lumped at the inlet, with k_c = 1.

		F1	F2	F3	F4	F5
	NT	502.2	1507.8	2521.7	3545.6	4580.4
	[ɑ]	642.3	1085.0	2469.0	3621.5	4134.5
	[e]	420.1	1973.7	2819.3	3650.6	4253.1
(A)	[i]	226.9	2276.1	3105.6	3728.7	4750.6
	[o]	505.2	867.7	2390.0	3457.7	4019.9
	[u]	237.2	600.0	2383.0	3710.5	4055.6
	[ɨ]	289.5	1529.4	2413.5	3471.1	4198.9
	NT	542.5	1518.0	2525.3	3549.4	4584.2
	[ɑ]	705.5	1108.6	2471.5	3625.1	4146.2
	[e]	464.8	1986.5	2824.1	3654.8	4264.3
(B)	[i]	301.6	2285.6	3118.3	3732.0	4756.1
	[o]	574.6	893.6	2392.7	3458.8	4023.6
	[u]	316.3	636.0	2384.8	3711.2	4057.9
	[ɨ]	357.0	1540.2	2419.7	3472.3	4200.4

Table VII. Reference formant frequencies for the Russian vowels.

The vocal tract is considered in closed conditions. Are taken into account the radiation impedance (Piston in Sphere model), and the viscous and thermal losses (shape factor = 1).

For the table (A) the wall impedance are supposed infinite.

For the table (B) area dependent lumped wall impedances are used.

NT	B1	B2	B3	B4	B5
(1)	4.8	8.7	11.4	13.8	15.9
(2)	2.8	29.5	80.1	142.1	205.3
(3)	10.6	1.0	.2	.1	.1
(4)	18.2	39.2	91.7	156.0	221.3
(5)	18.3	39.2	91.6	155.3	219.7
(6)	15.6	16.6	16.3	16.3	16.4
(7)	15.2	13.8	10.5	7.7	5.8

[i]	B1	B2	B3	B4	B5
(1)	4.8	12.6	27.0	19.6	25.8
(2)	.1	5.2	161.9	57.9	391.8
(3)	31.8	.6	.5	.1	.2
(4)	36.5	18.4	189.4	77.6	417.8
(5)	36.9	18.4	187.6	76.4	402.9
(6)	11.1	49.3	22.8	134.2	29.6
(7)	11.1	31.3	11.4	60.8	10.6

[a]					
(1)	8.4	11.3	18.1	18.8	28.2
(2)	5.2	15.6	40.2	134.5	28.3
(3)	12.3	3.4	.2	.2	.5
(4)	25.9	30.3	58.5	153.5	57.0
(5)	25.9	30.2	58.3	152.1	56.2
(6)	62.9	49.3	18.4	38.0	185.4
(7)	58.3	48.7	11.9	15.3	82.5

[o]					
(1)	6.5	9.6	14.6	23.2	22.5
(2)	2.7	7.9	15.5	7.3	20.1
(3)	15.6	4.8	.2	.0	.1
(4)	24.8	22.3	30.3	30.5	42.7
(5)	24.8	22.2	30.3	30.5	42.5
(6)	35.5	26.3	9.3	160.3	94.2
(7)	33.9	25.9	5.8	72.7	46.7

[e]					
(1)	5.0	11.5	15.3	17.6	19.0
(2)	1.2	29.0	106.8	327.9	283.8
(3)	13.5	1.0	.3	.2	.3
(4)	19.7	41.5	122.4	345.7	303.1
(5)	19.7	41.4	119.4	339.9	296.8
(6)	12.5	41.4	87.4	89.8	77.4
(7)	12.3	28.5	49.2	41.4	32.2

[u]					
(1)	6.1	9.1	19.2	20.2	21.1
(2)	.2	.2	.2	.4	1.2
(3)	30.4	9.4	.1	.0	.1
(4)	36.7	18.7	19.5	20.6	22.4
(5)	36.8	18.7	19.6	20.5	22.3
(6)	11.1	15.1	5.1	128.8	75.7
(7)	11.0	14.8	3.2	57.1	39.9

[ɨ]					
(1)	4.6	10.9	12.7	19.5	20.1
(2)	.3	64.8	10.2	32.5	31.0
(3)	25.9	1.0	.4	.1	.1
(4)	30.8	76.7	23.3	52.1	51.2
(5)	30.9	76.6	23.3	52.0	51.0
(6)	11.8	1.2	27.8	36.2	80.3
(7)	11.7	0.9	17.7	16.9	31.2

Table VIII Incremental bandwidths for the Russian vowels.

(1) Bandwidths for R & G losses,
(2) " for radiation losses,
(3) " for wall losses,
(4) Sum of lines (1), (2) and (3),
(5) Bandwidths for R & G, radiation and wall losses,
(6) " for glottal losses ($R_g = 270\Omega$ and $L_g = 0$),
(7) " for " " ($R_g = 270\Omega$ and $L_g = 16.67$ mH).

Remark : in all cases, the PIS model (imaginary part) and the wall area dependant impedances (imaginary part) were used. No subglottal system was used.

All the curves we give represent the transfer function between the pressure radiated from the mouth and the pressure source at the constriction. Whenever these specifications have been changed, it is explicitly mentioned.

Moreover, we have used, for the teeth pass, the internal end correction proposed by FANT (1960, p. 36):

$$l_i = 0.48(A_o)^{\frac{1}{2}} \cdot \left[1 - 1.25(A_o/A)^{\frac{1}{2}}\right], \qquad (52)$$

where l_i is the length correction for a tube of small area A_0 abruptly terminated in a wider tube of area A. Thus, the narrow tube corresponding to the teeth pass attains an effective length of 0.614 cm instead of 0.2 cm. The major consequence of this correction was the lowering of the third formant frequency by a few hundred Herz.

An example of a simulation is shown in Fig. 16, which illustrates the area function and associated input data and a calculated transfer function.

As a check of the reliability of the simulation, we made a computation with the same conditions as FANT (1960, p. 175) by forcing the glottis resistance to 5 ρc and the constriction resistance to 0.25 ρc. (see Fig. 17). The result checks fairly well with FANT's curve.

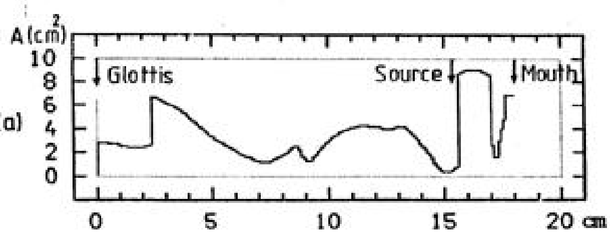
3.2.2 Influence of the glottis and constriction resistances

In order to study the influence of the glottis and the constriction resistances, we have made four simulations, see Fig. 18. The glottal area was $A_g = 0.10$ cm^2 and its inductance $L_g = 0$. Curve (1) corresponding to the zero constriction resistance and the infinite glottal resistance serves as a reference. From these curves it appears that the major effect of a finite, rather low, glottal resistance is to damp very strongly the first two formants, and the first zero, and also to damp and shift up in frequency the higher pole/zero pairs. The constriction resistance mainly damps F1 and F2 but also F3.

3.2.3 Influence of the glottal opening

The effect of the variation of glottal opening with a constant transglottal pressure $P_s = 8$ cmH_2O and with a glottal resistance and constriction resistance as dependent parameters is shown in Fig. 19. Part (a) corresponds to $L_g = 0$. It is verified that an opening of the glottis damps the first three formants and contributes to an overall spectral shape with only one broad peak, which is a characteristic of the [š] fricative. Part (b) shows that the glottal inductance does not play an important role.

(a)

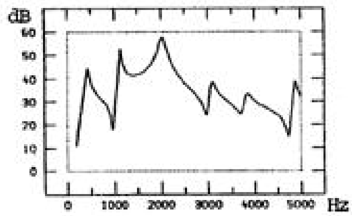


(b)

```
** VCTR00 **

Minimum frequency (Hz)     : 200
Maximum frequency (Hz)     : 5000
Frequency step (Hz)        : 20
R & C losses ? (Shape Factor) : 1
Characteristic Impedance (0 = Real, 1 = Complex) : 0
Radiation System ?
  (0 = Piston in Sphere,
   1 = 0.8*R0 Lengthening (with tg),
   2 = Wakita-Fant (0.8*R0 Lengthening),
   3 = Free Space,
   4 = Stevens-Kasowski-Fant (Analog Simulation)     : 0
Radiation Impedance ?
  (0 = Complex, 1 = Imaginary, 2 = Real) : 0
Wall Impedance (0 = Infinite, 1 = Complex, 2 = Imaginary,
 3 = Ce method (Complex), 4 = Ce method (Imaginary)) : 3
Glottis Area (cm2)        : 0.025
Transglottal pressure   (cmH20) : 8
Glottis length (cm)        : 1.2
Glottis thickness (cm)    : 0.3
Glottis impedance ?
  (0 = Complex, 1 = Imaginary, 2 = Real, 3 = Infinite) : 1
Subglottal system (0-1) : 0
Shape factor at the inlet of the glottis : 1.
Constriction resistance ? (0-1) : 1
Shape factor at the inlet of the constriction : 1.
```

(c)



(d)

```
***  Vocal Tract Simulation  ***                    (25/10/84)

                    SH.HN

Minimum frequency =   200.
Maximum frequency =  5000.
Frequency step    =  20.00

        Characteristics of the Vocal Tract :

R & C losses (Shape Factor = 1.0)
Characteristic Impedance : Real
Radiation System : Piston in Sphere
Radiation Impedance : Complex
Radiation transfer (+6dB/Oct)
Wall Impedance (Ce method (Complex))

        Characteristics of the glottis :

Glottis area =  .025 cm2
Glottis length = 1.20 cm
Glottis thickness = .30 cm
Transglottal pressure =  8.0 cmH20
Shape factor at the inlet of the glottis = 1.00
Glottis resistance =  61.0468 + 169.2026 = 230.2493 Ohms
Glottis inductance =  .000000 H

        Characteristics of the constriction :

Constriction geometry : L =  .40 cm , D =  .71 cm
Shape factor at the inlet of the constriction = 1.00
Constriction resistance =    .6609 +    .0104 =    .6713 Ohms
REYNOLDs number at the constriction =  1025.

        Poles, Zeros and Bandwidths values :

  Pole No 1 : F =   417.0, B =   64.8, Amp =   44.2, corr. =   -.0
  Zero No 1 : F =   967.1, B =   40.1, Amp =   17.5, corr. =   -.5
  Pole No 2 : F =  1116.8, B =   42.4, Amp =   52.5, corr. =  -1.1
  Pole No 3 : F =  2021.5, B =  119.5, Amp =   57.8, corr. =  -5.0
  Zero No 2 : F =  2982.5, B =   73.9, Amp =   24.3, corr. =   5.3
  Pole No 4 : F =  3070.5, B =   93.9, Amp =   38.0, corr. =  -1.7
  Zero No 3 : F =  3727.3, B =  142.5, Amp =   25.6, corr. =  -2.0
  Pole No 5 : F =  3806.7, B =  137.7, Amp =   32.3, corr. = -11.9
  Zero No 4 : F =  4744.9, B =   69.6, Amp =   15.1, corr. =  -2.2
  Pole No 6 : F =  4876.2, B =   69.6, Amp =   38.2, corr. =   -.2
```

Fig. 16. Example for input and output data for a simulation.
(a) graphic representation of the vocal tract area function,
(b) keyboard input,
(c) graphic representation of the transfer function between pressure radiated at the mouth and the consonantal pressure source.
(d) poles and zeros corresponding to the previous curve.

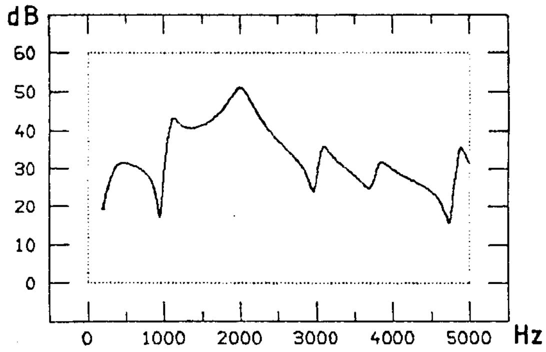


Fig. 17. Transfer function between radiated pressure at the mouth and consonantal source pressure, for the /š/ configuration in the FANT (1960) conditions, that is:
- infinite wall impedances,
- glottis resistance of 5 ρc (no glottis inductance),
- tongue pass resistance of 0.25 ρc,
- pressure source in the middle of the constriction.

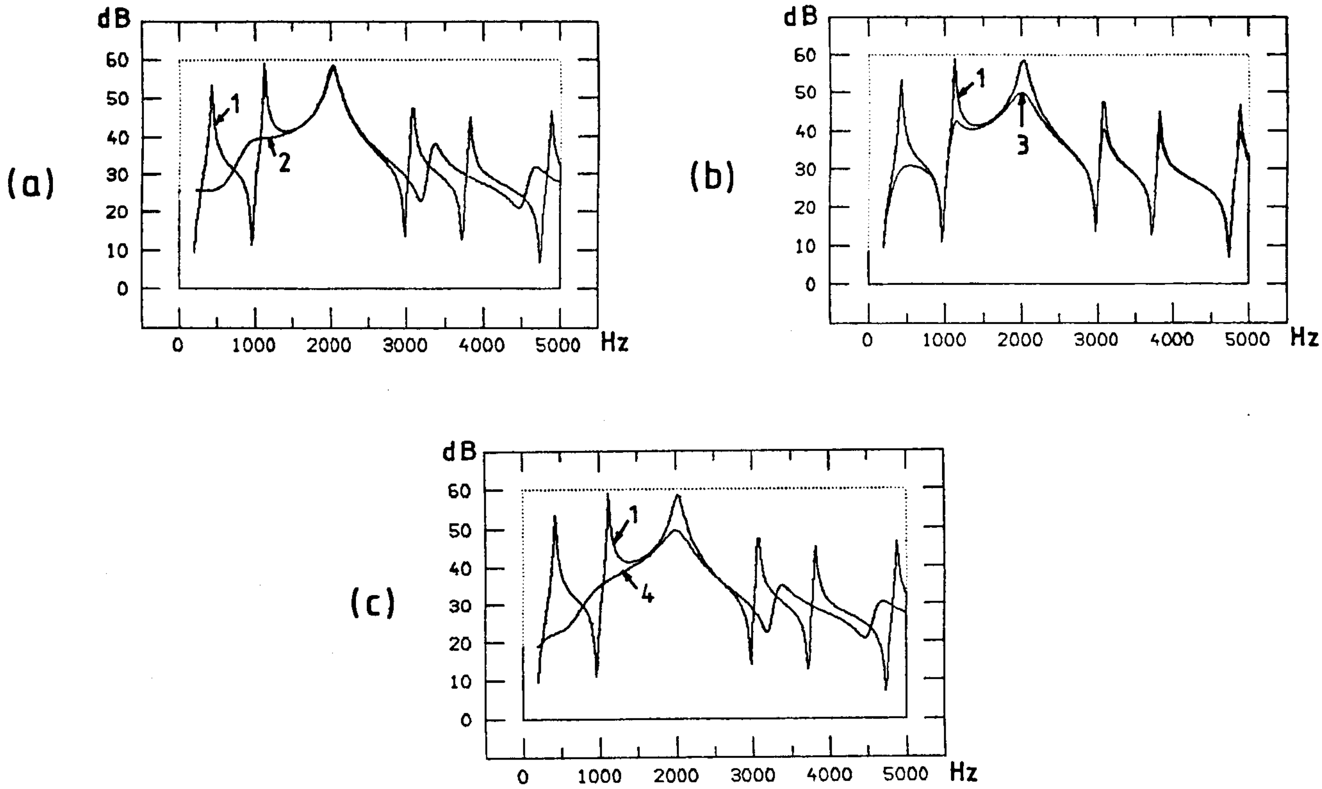


Fig. 18. Influence of the glottis and constriction resistances.
(1) neither glottis nor constriction resistances (L_g=0),
(2) glottis resistance only (R_g=8.5, L_g=0),
(3) constriction resistance only (R_c=13.3, L_g=0),
(4) glottis and constriction resistances (L_g=0).

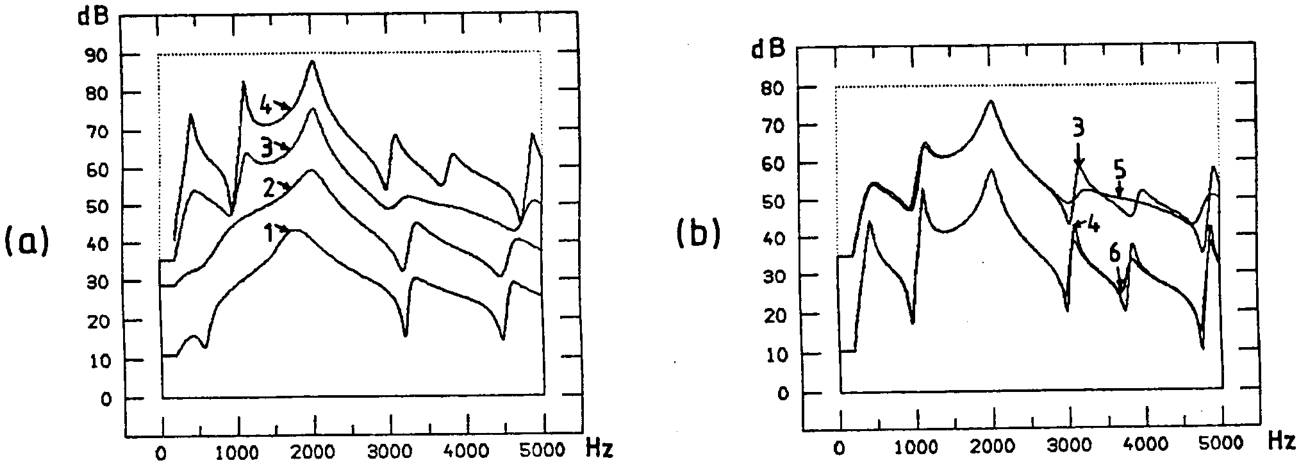


Fig. 19. Influence of the glottal opening.

(a) no glottal inductance:

(1) $A_g = 2.0\ cm^2$; $R_g = 2.1\Omega$; $R_c = 53.1\ \Omega$;
(2) $A_g = 0.5\ cm^2$; $R_g = 8.5\Omega$; $R_c = 13.3\ \Omega$;
(3) $A_g = 0.1\ cm^2$; $R_g = 43.2\ \Omega$; $R_c = 2.67\Omega$;
(4) $A_g = 0.025\ cm^2$; $R_g = 230.0\ \Omega$: $R_c = 0.67\Omega$;

(the four curves are shifted 10 dB from each other for better readability)

(b) with glottal inductance:

(3) and (4) are the same as in (a), without glottal inductance,
(5) $A_g = 0.1\ cm^2$; $R_g = 43.2\Omega$; $L_g = 3.4$ mH; $R_c = 2.67\Omega$;
(6) $A_g = 0.025 cm^2$; $R_g = 230.0\Omega$; $L_g = 13.7$ mH; $R_c = 0.67\Omega$;

(the curves (3) and (5) are shifted 20 dB above (4) and (6)).

3.2.4 Influence of the source location

Fig. 20 shows the influence of the source location on the transfer function. It appears that, as long as the source is located in the vicinity of the constriction where the corresponding area is small, a small shift of the source has an effect limited to a small shift in the frequency and the level of high frequency zeros. But if the source is located inside the large cavity downstream of the constriction, a small shift of the source leads to a substantial shift in frequency of the upper zeros and some reduction in the F3 level. These zero shifts were expected since the volume of the part of the front cavity behind the constriction varies significantly when the source is shifted due to the fairly large cross-sectional area. When the source is shifted to the inlet of the teeth path (curve (6)), the main effect is a high frequency emphasis related to an additional zero below the main formant.

We may conclude from this preliminary study, that our model is a useful tool allowing a detailed study of different aspects of speech production theory, including consonants.

3.3 HELIUM SPEECH

In this section we compare the poles and bandwidths obtained for normal speaking conditions, that is, the air at a pressure of 1 Atmosphere, with those obtained for a presumed diving condition assuming a depth of 300 m and a helium/oxygen breathing mixture of 99% He and 1% O_2.

For both conditions studies have been undertaken of incremental bandwidths, i.e., differential contributions from various dissipative elements within a neutral tube model. For the set of Russian vowels, the comparison involves formant frequencies and total bandwidths.

The glottis state is the one used in Section 2.4.2 (A_g = 0.027 mm^2, P_s = 16 cm H_2O, l_g = 1.2 cm, d = 0.3cm, k_g = 1) for all the calculations. The following physical constants have been adopted:

$$\left\{\begin{aligned} c_h &= c_o \times 2.79 \\ \rho_h &= \rho_o \times 4.24 \\ C_{ph} &= C_{po} \times 5.09 \\ (\eta-1)_h &= (\eta-1)_o \times 1.57 \\ \mu_h &= \mu_o \times 1.11 \\ \lambda_h &= \lambda_o \times 5.01 \end{aligned}\right. \qquad (53)$$

the subscript h standing for heliox at 300m depth, and the subscript 0 for air at normal pressure*.

---

* These have been chosen to allow comparisons with results from calculations of M.R. RICHARDS & R.W. SHAFER (personal communication).

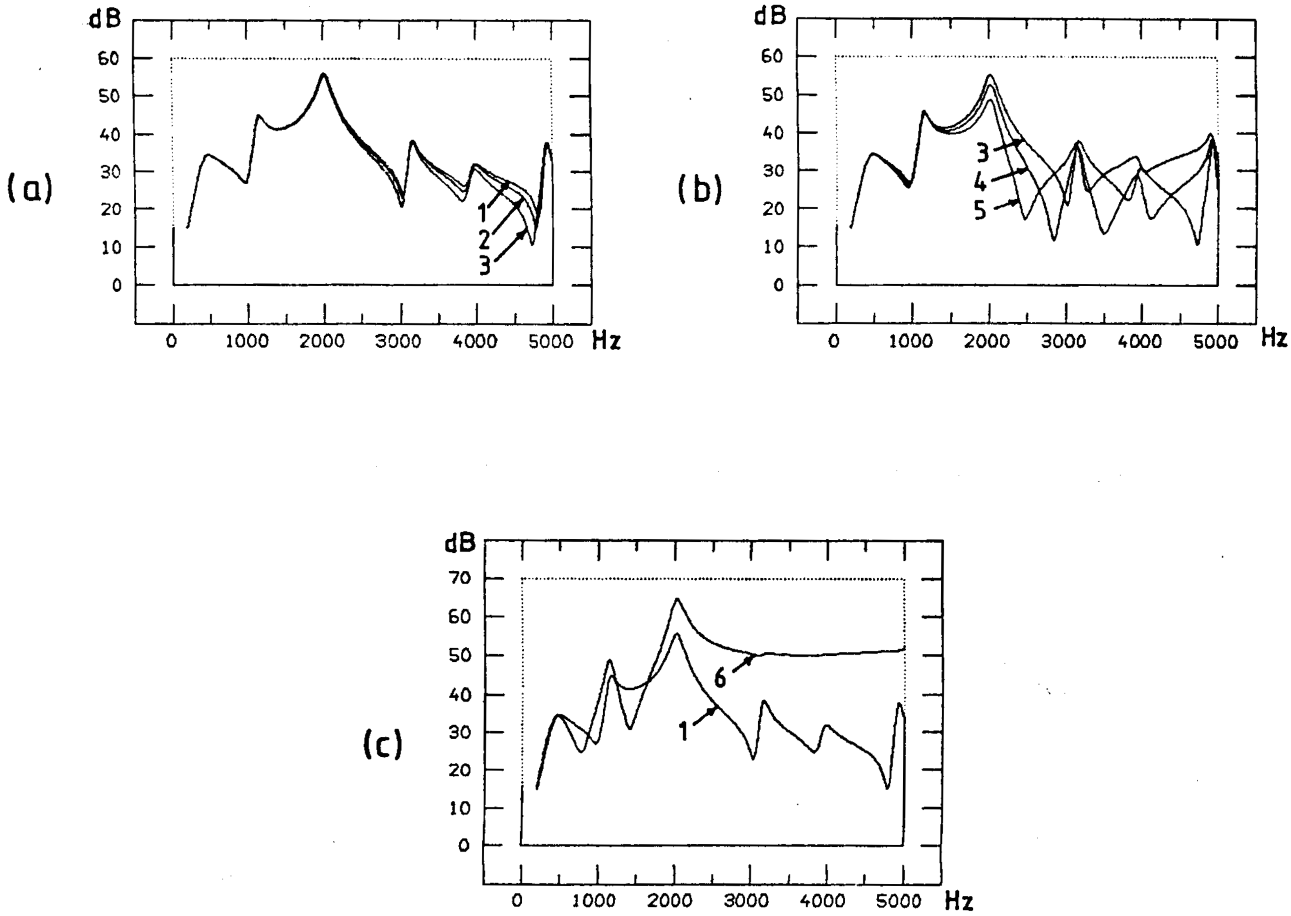


Fig. 20. Influence of source location.
The figure has been split into three parts for better readability. $x_s = 0$ corresponds to a source located at the outlet of the constriction, which is the tube of 0.4 cm length and 0.4 cm^2 area (see Fig. 16a).

(1) -> x_s = −0.4 cm, (4) -> x_s = +0.4 cm,
(2) -> x_s = 0.0 cm, (5) -> x_s = +0.5 cm,
(3) -> x_s = +0.3 cm, (6) -> dental source.

The conditions are $A_g = 0.1\ cm^2$, $R_g = 43.2\Omega$, $L_g = 3.4$ mH, $R_c = 2.67\,\Omega$.

3.3.1 Incremental bandwidths for the uniform tube

In the loss-less case, without the wall loading and radiation impedances, the poles frequencies are evidently proportional to the velocity of the sound, c, and, thus, would be expected to increase by the factor $K = c_h/c_0$. Taking into account the wall impedance loading (FANT, 1972; FANT & LINDQVIST, 1968), by means of Eq. 38a, we arrive at a frequency shift factor

$$K = \frac{C_h}{C_o}\left[1 + \frac{F_{wo}^2}{F_{wh}^2}\left(\frac{\rho_h}{\rho_o} - 1\right)\right]^{\frac{1}{2}} \tag{54}$$

The expected shifts of incremental bandwidths will next be treated.

The viscous and thermal losses are subjected to the following proportionalities:

$$B_R \sim C\cdot\alpha_R \sim \sqrt{\frac{\mu}{\rho}}\,\sqrt{f}; \text{ correcting factor: } 0.51\sqrt{f_h/f_o} \tag{55a}$$

$$B_G \sim C\cdot\alpha_G \sim (\eta-1)\sqrt{\frac{\lambda}{C_p\rho}}\,\sqrt{f}\;; \quad \text{correcting factor: } 0.82\sqrt{f_h/f_o} \tag{55b}$$

Since B_R corresponds to 69% of the total bandwidth and B_G to 31%, the expected correcting factor is:

$$(0.51 * 0.69 + 0.82 * 0.31)\sqrt{f_h/f_o} = 0.6\sqrt{f_h/f_o} \tag{56}$$

If R_0 is the radiation resistance, and if we neglect the wall impedances and the radiation inductance, we can show that the corresponding bandwidth is given by:

$$B = \frac{c}{\pi l}\,\frac{R_o}{Z_o} \sim \frac{f^2}{c}\,, \tag{57}$$

if we suppose $K_s(\omega) = 1$.

For the wall losses, we can derive:

$$B_w = \frac{B_{wo}\cdot F_w^2}{f^2 + B_{wo}^2} \simeq \frac{B_{wo}\cdot F_w^2}{f^2} \tag{58}$$

From FANT (1972, p. 29), we know that:

$$F_w \sim c\sqrt{\rho}, \tag{59}$$

and so

$$B_w \sim F_w{}^2/f^2 \sim c^2{}_\rho/f^2 \ . \tag{60}$$

For the glottis resistance, if we neglect all the inductances (walls and radiation), we can show that the corresponding bandwidth is:

$$B = \frac{c}{\pi l} \ \frac{Z_o}{R_g} \sim c \ \frac{\rho c}{\sqrt{l}} = c^2\sqrt{\rho} \ . \tag{61}$$

The results are given in the Tables IXa and IXb for two different conditions : without glottis inductance and with glottis inductance. One can verify that they check, at least in order of magnitude. Some discrepancies may be due to the fact that these formulas are approximations only.

In the heliox pressure case, the glottal inductance L_g being proportional to ρ, dominates the glottal impedance and, thus, screens off the glottal resistance from the vocal tract. Accordingly, glottal losses are minimized. Without glottal inductance, on the other hand, the glottal bandwidths increase by a factor of the order of 16. The extreme glottal bandwidths have relevance for "leaky" voices only.

3.3.2 Russian vowels-Heliox conditions

In Table X the pole frequencies and bandwidths for the Russian vowels are compared for normal speaking conditions and for the heliox mixture.

As expected the frequency shift is very close to c_h/c_0 for F3, F4, and F5. As a consequence of the wall impedances, the ratio is much higher for F1 and F2 and is approximately predictable from Eq. 38a. The bandwidth shifts display more complex patterns. However, a general observation is that the order of magnitude of bandwidths shift factors is c_h/c_0. The wall impedances have a small part only in this shift and glottal losses are limited by the assumption of a finite glottis inductance.

Appendix

Here we give the values used throughout the paper for the different physical constants. The values for air are given at normal atmospheric pressure.

Normal conditions

Form. No	1	2	3	4	5
Bandwidths					
R & G	4.8	8.7	11.4	13.8	15.9
Rad.	2.8	29.5	80.1	142.1	205.3
Walls	10.6	1.0	.2	.1	.1
Glot.	15.6	16.6	16.3	16.3	16.4
All	33.9	55.6	107.5	170.6	234.3
Freq.	542.6	1517.8	2525.0	3549.3	4584.2

Heliox mixture

Form. No	1		2		3		4		5	
Bandwidths										
R & G	4.7	0.98	8.7	1.00	11.6	1.02	13.9	1.00	16.1	1.01
Rad.	9.5	3.39	77.9	2.64	219.7	2.74	402.8	2.83	587.2	2.86
Walls	28.4	2.68	4.3	4.30	.9	4.50	.5	5.00	.6	6.00
Glot.	249.8	16.01	321.4	19.36	294.2	18.05	273.7	16.79	274.5	16.74
All	292.0	8.61	408.8	7.35	515.9	4.80	669.7	3.92	843.5	3.60
Freq.	1826.1	3.37	4326.2	2.85	7072.5	2.80	9928.9	2.80	12834.6	2.80

Table IXa. omparison between the incremental bandwidths for normal speaking conditions and for heliox mixture, without glottis inductance, for a uniform tube (16.54 cm, 6 cm^2).

In all cases, the imaginary part of the radiation load (PIS model) and of the wall impedances (area independant lumped impedances) were used.

In the second part of the table, for each formant, the second nummer is the ratio between the value for the heliox mixture and the one corresponding to the normal conditions.

The glottis resistance is 270 Ω for the normal conditions and 510 for the heliox mixture.

Normal conditions

Form. No	1	2	3	4	5
Bandwidths					
R & G	4.7	8.6	11.4	13.8	15.9
Rad.	3.1	29.8	80.5	142.0	205.0
Walls	8.9	1.1	.2	.1	.1
Glot.	15.2	13.8	10.5	7.7	5.8
All	33.5	52.9	102.1	162.7	224.8
Freq.	543.8	1521.0	2529.0	3553.2	4587.8

Heliox mixture

Form. No	1		2		3		4		5	
Bandwidths										
R & G	4.7	1.00	8.7	1.01	11.5	1.01	13.9	1.01	16.0	1.01
Rad.	10.6	3.42	78.9	2.65	220.3	2.74	403.2	2.84	587.6	2.87
Walls	24.8	2.79	4.4	4.00	.9	4.50	.5	5.00	.6	6.00
Glot.	93.8	6.17	33.7	2.44	12.7	1.21	6.1	0.79	3.7	0.64
All	135.6	4.05	125.7	2.38	245.1	2.40	422.8	2.60	606.1	2.70
Freq.	1881.6	3.46	4378.8	2.88	7108.4	2.81	9950.0	2.80	12842.2	2.80

Table IXb. Comparison between the incremental bandwidths for normal speaking conditions and for heliox mixture, with glottis inductance, for a uniform tube (16.54 cm, 6 cm^2).

In all cases, the imaginary part of the radiation load (PIS model) and of the wall impedances (area independant lumped impedances) were used.

In the second part of the table, for each formant, the second nummer is the ratio between the value for the heliox mixture and the one corresponding to the normal conditions.

The glottis resistance and inductance are 270Ω and 12.67 mH for the normal conditions and 510Ω and 53.7 mH for the heliox mixture.

	(A)	(B)	(C)	(D)	(E)	(F)	(G)
	543.8	1881.6	3.46	3.30	33.5	135.6	4.05
	1521.0	4378.8	2.88	2.86	52.9	125.7	2.38
NT	2529.0	7108.4	2.81	2.82	102.1	245.1	2.40
	3553.2	9950.0	2.80	2.80	162.7	422.8	2.60
	4587.8	12842.2	2.80	2.80	224.8	606.1	2.70
	716.3	2463.8	3.44	3.09	82.8	146.8	1.77
	1113.0	3633.9	3.26	2.92	72.8	277.4	3.81
[ɑ]	2475.8	6954.2	2.81	2.82	70.2	144.8	2.06
	3633.1	10181.4	2.80	2.80	165.0	418.8	2.54
	4193.3	11900.2	2.84	2.80	128.1	150.6	1.18
	465.8	1671.3	3.59	3.46	32.1	119.3	3.72
	1996.2	5738.8	2.87	2.83	69.6	139.4	2.00
[e]	2845.4	8068.6	2.84	2.81	168.4	356.5	2.12
	3675.2	10362.1	2.82	2.80	373.4	995.5	2.67
	4297.3	12060.5	2.81	2.80	320.2	830.0	2.59
	302.0	1355.0	4.49	4.22	48.1	155.6	3.23
	2297.9	6555.0	2.86	2.82	49.1	68.9	1.40
[i]	3122.2	8824.9	2.83	2.81	197.3	471.8	2.39
	3766.1	10609.8	2.82	2.80	132.9	228.0	1.72
	4767.5	13391.0	2.81	2.80	412.0	1156.0	2.81
	580.0	2050.7	3.54	3.24	58.1	105.3	1.81
	894.6	2969.0	3.32	2.99	48.1	216.2	4.49
[o]	2395.0	6718.4	2.81	2.82	36.1	66.3	1.84
	3505.2	9861.9	2.81	2.80	108.3	106.0	0.98
	4047.5	11385.5	2.81	2.80	88.3	112.5	1.27
	317.5	1294.7	4.08	4.10	48.0	84.4	1.76
	636.3	2220.0	3.49	3.17	33.5	150.3	4.49
[u]	2386.0	6678.6	2.80	2.82	22.7	24.3	1.07
	3751.5	10519.4	2.80	2.80	76.8	59.9	0.78
	4077.1	11444.1	2.81	2.80	57.1	53.6	0.94
	357.6	1498.6	4.19	3.84	42.8	148.6	3.47
	1540.4	4395.4	2.85	2.86	77.5	194.9	2.51
[ɨ]	2426.4	6858.9	2.83	2.82	41.0	65.0	1.59
	3481.7	9745.7	2.80	2.80	70.3	128.4	1.83
	4220.2	11818.1	2.80	2.80	81.8	130.3	1.59

Table X. Frequency and bandwidth changes, for the Russian vowels, between normal speaking conditions and heliox mixture.

(A) pole frequencies for normal conditions,
(B) " " " heliox mixture,
(C) ratio (B)/(A),
(D) predicted ratio assuming F_{WO} = 190 Hz,
(E) bandwidths for normal conditions,
(F) " " heliox mixture,
(G) ratio (E)/(D).

All losses were used (including PIS radiation impedance and area dependant lumped wall impedances). The glottis conditions are the same as for the tables 7. There was no subglottal system.

Sound velocity	: $c = 35300\ cm.sec^{-1}$,
Density of air	: $\rho = 1.14\ 10^{-3}\ g.cm^{-3}$,
Specific heat at constant pressure	: $C_p = 2.4\ 10^{-1}\ cal.g^{-1}.deg^{-1}$,
Adiabatic gas constant	: $\eta = 1.4$,
Viscosity of air	: $\mu = 1.84\ 10^{-4}\ dyne.sec.cm^{-2}$,
Coefficient of heat conduction of air	: $\lambda = 5.5\ 10^{-5}\ cal.cm^{-1}.sec.deg^{-1}$.

References

ANANTHAPADMANABHA T.V. & FANT G. (1982): Calculation of the true glottal flow and its components, Speech Communication 1, pp. 167-184.

ATAL B.S., CHANG J.J., MATHEWS M.W., & TUCKEY J.W. (1978): "Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique", J.Acoust.Soc.Am. 63, pp. 1535-1555.

van den BERG Jw. (1953): "Physica van de stem vorming met toepassingen", Diss. Rijksuniversiteit te Groningen ('s-Gravenhage).

van den BERG Jw., ZANTEMA J.T., & DOORNENBAL P. (1957): "On the air resistance and the Bernouilli effect of the human larynx", J.Acoust.Soc. Am. 29, pp. 626-631.

FANT G. (1960): Acoustic Theory of Speech Production, Mouton, ('s-Gravenhague, 2nd edition, 1970).

FANT G. (1969): "Stops in CV-syllables", STL-QPSR 4/1969, pp. 1-25.

FANT G. (1972): "Vocal tract wall effects, losses, and resonance bandwidths", STL-QPSR 2-3/1972, pp. 28-52.

FANT G. (1985): "The vocal tract in your pocket calculator", To appear in Linguistic Phonetics, ed. by Victoria FROMKIN.

FANT G. & ANANTHAPADMANABHA T.V. (1982): "Truncation and superposition", STL-QPSR 2-3/1982, pp. 1-17.

FANT G. & LILJENCRANTS J. (1979): "Perception of vowels with truncated intraperiod decay envelopes", STL-QPSR 1/1979, pp. 79-84.

FANT G. & LINDQVIST J. (1968): "Pressure and gas mixure effects on diver's speech", STL-QPSR 1/1968, pp. 7-17.

FANT, G. & SONESSON, B. (1964): "Speech at high ambient air-pressure", STL-QPSR 2/1964, pp. 9-21.

FANT G., ISHIZAKA K., LINDQVIST-(GAUFFIN) J., & SUNDBERG J. (1972): "Subglottal formants". STL-QPSR 1/1972, pp. 1-12.

FANT G., NORD L., & BRANDERUD P. (1976): "A note on the vocal tract wall impedance", STL-QPSR 4/1976, pp. 13-20.

FLANAGAN J.L. (1972): Speech Analysis, Synthesis and Perception, Springer Verlag Berlin, Heidelberg, New-York.

ISHIZAKA K., FRENCH J.C., & FLANAGAN J.L. (1975): "Direct determination of vocal tract wall impedance", I.E.E.E. Trans. A.S.S.P. 23, pp. 370-373.

LILJENCRANTS J. & FANT G. (1975): "Computer program for VT-resonance frequency calculation", STL-QPSR 4/1975, pp. 15-20.

MORSE P.M. (1976): Vibration and sound, Ed. by J.Acoust.Soc.America.

MRAYATI M. (1976): "Contribution aux études sur la production de la parole", "Modèles électriques du conduit vocal avec pertes, du conduit nasal et de la source vocale", "Etude de leurs interactions", "Relations entre dispositions articulatoires et caractéristiques acoustiques", Thèse d'Etat, I.N.P., Grenoble.

PORTNOFF M. (1973): "A quasi one dimensional digital simulation for the time-varying vocal tract", M.S. Thesis, M.I.T.

STEVENS K.N., KASOWSKI S., & FANT G. (1953): "An electrical analog of the vocal tract", J.Acoust.Soc.Am. 25, pp. 734-742.

WAKITA H.& FANT G. (1978): "Toward a better vocal tract model", STL-QPSR 1/1978, pp. 9-29.