BRITISH ACOUSTICAL SOCIETY

"SPRING MEETING" at Chelsea College, London, S.W.3 on Wednesday 25th April / Friday 27th April, 1973.

SPEECH AND HEARING: Session 'A': Speech Production and Perception.

Paper No. 73SHA3

Vocal Tract Damping and Cavity Wall Vibrations

Gunnar FANT, Department of Speech Communication
Royal Institute of Technology (KTH)
S-100 44 Stockholm 70, Sweden

(Invited paper)

Summary

This report summarizes results of a study published in STL-QPSR 2-3/1972 the purpose of which was to extend the theoretical basis for calculating the damping of vocal tract resonance modes and to match this theory with empirical data from the Fujimura-Lindqvist vocal tract sinewave sweep studies. In general, cavity wall losses dominate at low frequencies, friction and heat conduction losses at middle frequencies, and radiation losses at higher frequencies. Important deviations from average trends exist especially with respect to the effects of radiation which differ appreciably depending on the particular vocal cavity configuration. Surface losses may be taken into account by overall vocal tract parameters and the wall losses are fairly stable. One outcome of the study is that empirical formulas have been constructed for predicting each resonant bandwidth from the entire set of VT resonance frequencies.

The sound propagated through the cavity walls and radiated externally from the neck adds a finite low frequency correction to the sound radiated from the lips. This effect is enhanced in speech under hyperbaric conditions which also provides favorable experimental means for estimating cavity wall impedance.

Cavity wall impedance

Under the closed glottis conditions the dissipative sources within the vocal tract are the radiation resistance R_ρ and the distributed series elements $R(x)$ representing frictional losses and the parallel elements $G(x)$ representing the heat conduction losses. In parallel to the latter enters the distributed cavity wall resistance $R_W(x)$ in series with an associated wall inductance $L_W(x)$. With the lip end closed the resonance frequency of the tract is determined by the total lumped wall inductance and the volume of the enclosed air. This limiting F_1 value of the closed tract is denoted F_W and is of the order of 175 Hz for a male voice. For the Helmholtz resonator case it is apparent that

$$F^2 = F_i^2 + F_W^2 \tag{1}$$

where F_i is the ideal resonance frequency without mass shunt. It has been shown that Eq. (1) also holds for the single open tube with even distribution of mass loading providing

$$F_W^2 = 1/4\pi^2 L(x)C(x) \tag{2}$$

and for an arbitrarily shaped vocal tract, providing the thickness of the cavity walls d(x) is proportional to s(x)/A(x), the ratio of cross-sectional perimeter to cross-sectional area, in which cases Eq. (1) has a general validity for any resonance.

Under hyperbaric conditions $F_W^2 = (P/P_o)F_{Wo}^2$, where P_o is 1 atmosphere and P the static pressure. This is a situation favorable for observing F_W directly or from observations of shifts in F_1 with P through Eq. (1). The F_W values thus derived do not appear to vary much with the particular vowel or speaker and were found to lie within the range 150-200 Hz and accounts for the typical F^{-2} frequency dependency of B1 towards the lower end of the F_1-range. Under the conditions stated above the resonance bandwidth contribution from wall losses is

$$B_W = B_{Wo}\left(\frac{F}{F}W\right)^2 \qquad (3)$$

where the closed tract resonance bandwidth B_{Wo} is of the order of 100 Hz for a male voice and about 20 % higher for an average female voice. This is the main source of F1 damping at frequencies below 350 Hz.

Surface losses

The bandwidth terms B_R and B_G associated with R(x) and G(x) increase as $A(x)^{-1/2}$ and as $F^{1/2}$ for tube modes and as $F^{-1/2}$ for Helmholtz resonator modes, e.g. for a first formant tuned by a varying mouth opening area at constant vocal tract volume. The

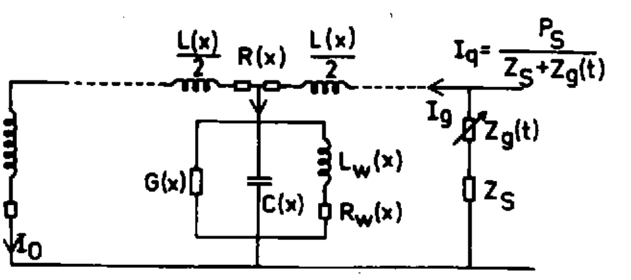


Fig. 1

articulatory narrowing of a vowel thus tunes F_1 to a lower frequency and greater B_1 adding to the inverse frequency-bandwidth relation of Eq. (3). One outcome of the Fant (1972) study is that an approximation to B_R+B_G may be obtained by substituting any arbitrary area function by a single tube of the same mean area determined by integrating 1/A(x) over the tract. The difference was maximally 6 Hz for any of B_1, B_2, and B_3 in 5 vowels studied.

Radiation losses

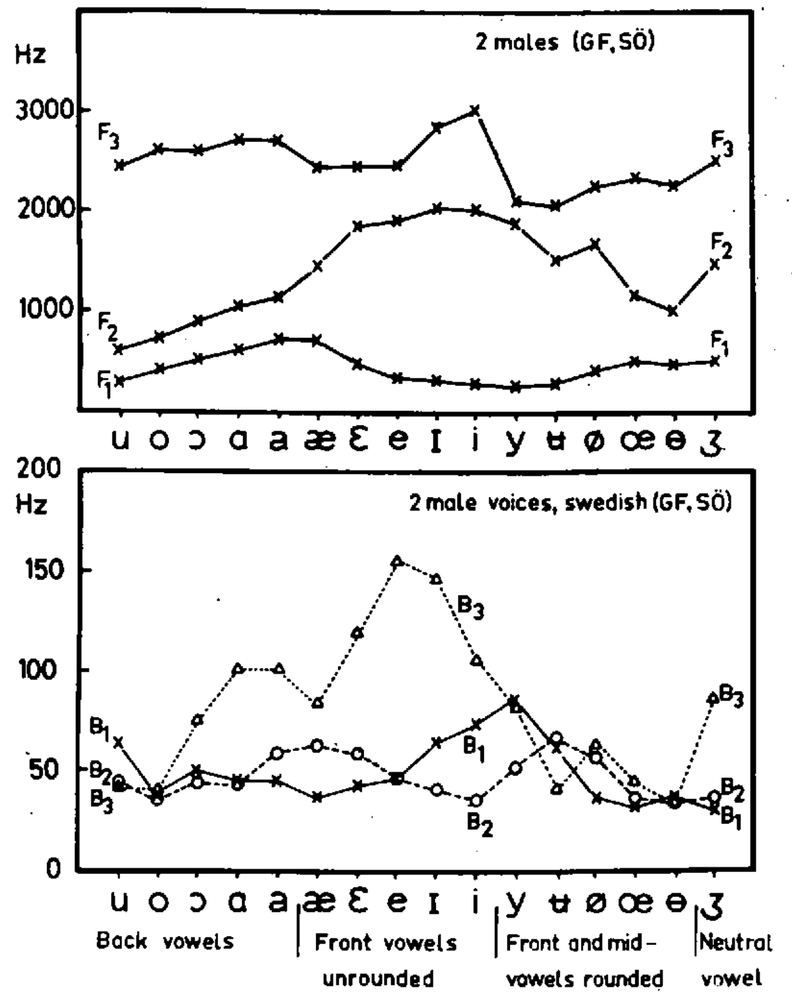
Since the radiation component B_o of resonance bandwidths may attain quite large values at frequencies above 2000 Hz it pays to carry out calculations with a reasonable estimate of the particular baffle effect. The infinite baffle has twice the radiation resistance of a point source and a representative intermediate value for speech is 1.4 times the point source resistance. In the open tube resonator B_o is proportional to F^2, A, and l_e^{-1}. The single tube formula is useless for estimating the radiation damping of standing waves in the pharynx such as F2 of the vowel [i] which is effectively decoupled from the radiation resistance and is influenced by this element to an extremely small fraction only. The main sources of dissipation are those inside the cavity, i.e. in conformity with a general rule that dissipation of energy is proportional to the stored kinetic and potential energies per unit length and the resistive elements of type R(x) respectively G(x).

As the point of constriction in the tract is shifted the radiation damping of a mode displays large variations depending on the particular mode-cavity dependencies.

Empirical data

To what extent may bandwidths be predicted from the total set of resonance frequencies F_1 F_2 F_3 F_4 of a particular sound? The evidence is brought out in Fig. 2 which is a replotting of data from Fujimura-Lindqvist. The high B_1 of low F_1 vowels is apparent.

Fig. 2.



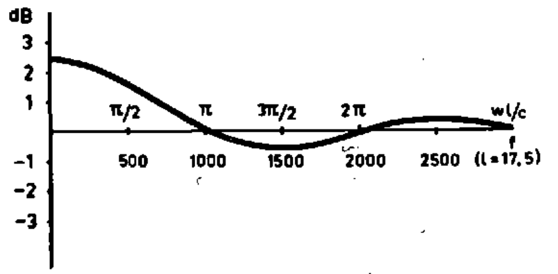
It is seen that a relation $B_1 < B_2 < B_3$ exists only for the non high front vowels whereas in [u] $B_1 > B_2 > B_3$. All rounded vowels have small B_3. The extreme range of variation with articulation is as follows, $29 < B_1 < 87$, $34 < B_2 < 75$, $34 < B_3 < 155$. An attempt has been made to describe the empirically found correlations between mode frequency patterns and bandwidths as follows.

$$\left.\begin{aligned} B_1 &= 15(500/F_1)^2 + 20(F_1/500)^{1/2} + 5(F_1/500)^2 \\ B_2 &= 22 + 16(F_1/500)^2 + 12000/(F_3-F_2) \\ B_3 &= 25(F_1/500)^2 + 4(F_2/500)^2 + 10F_3(F_{4a}-F_3) \end{aligned}\right\} \quad (4)$$

where F_{4a} is the average F_4 of the speaker. The average deviation between empirical measures and the predicted values is of the order of 1 decibel for B_1 and B_2 and 2 decibels for B_3 of this male group. The female group gave even better correlations except for the [u:] vowel.

In normal voiced speech the finite glottal resistance and inductance should provide additional contributions ranging from 0-50 Hz in the maximally open phase of the glottal cycle. The magnitude of the glottal damping is just as mode-specific as was discussed for radiation and largely confined to F_1 and typical back-cavity formants. The flow induced increase in frictional losses can be appreciable for voiced fricatives, see Fant (1960).

Fig. 3.



Superposition effects

The sound propagated through the vocal tract walls and radiated at the external surfaces accounts for a frequency correction in the overall response as exemplified by Fig. 3 which pertains to a lossless lumped mass introduced at the level of the glottal source in a uniform tube model of the vocal tract. The initial amplitude of the sinx/x function which is 2.5 dB increases proportionally to the atmospheric pressure and may reach high values in hyperbaric speech and zeros may become apparent. In normal speech a similar mechanism accounts for some degree of redistribution of zeros in the spectra of nasal consonants.

References

Fant, G. (1960): Acoustic Theory of Speech Production ('s-Gravenhage; 2nd edition 1970).

Fant, G. (1972): "Vocal Tract Wall Effects, Losses, and Resonance Bandwidths", STL-QPSR 2-3/1972 (KTH, Stockholm), pp. 28-52.

Fujimura, O. and Lindqvist, J. (1971): "Sweep-Tone Measurements of Vocal-Tract Characteristics", J. Acoust. Soc. Am. 49 (1971), pp. 541-558.