TR－H－108

A new method for measuring vocal tract transmission characteristics

Jianwu Dang
Kiyoshi Honda

1994. 11. 22

ATR人間情報通信研究所

〒619-02 京都府相楽郡精華町光台2-2　☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011
Facsimile: +81-7749-5-1008

A new method for measuring vocal tract transmission characteristics

Jianwu DANG and Kiyoshi HONDA

ATR Human Information Processing Research Labs.,

2-2 Hikaridai Seikacho Soraku-gun, Kyoto, 619-02 Japan

This paper has been submitted for publication in the Journal of the Acoustical Society of America.

ABSTRACT

A new method is proposed for the purpose of direct measurement of vocal tract transmission characteristics. This method evaluates the vocal tract transfer function from an arbitrary location in the vocal tract to its radiation end by measuring volume velocities at the two locations. An experimental study was conducted on acoustic tubes of various known shapes in order to evaluate the accuracy of the method. Two types of excitation sound sources, a sweeptone signal and the natural glottal sound source, were used for the experiment. Resonance and antiresonance frequencies obtained from the acoustic tubes were consistent with theoretical values within 4%. Compared with previous methods (Van den berg, 1955; Fujimura, et al, 1971; Lindqvist-Gauffin, et al, 1976), the present method has two major advantages: first, the characteristics and positions of the excitation sound source have less influence on the measurement of the frequencies of both the peaks and zeros; second, the measured results exactly represent acoustic properties of the measurement portion from the measurement location to the radiation end. By taking these advantages, the glottal sound can serve as the excitation sound source for the direct measurement. The spectrum details of the vocal tract transmission characteristics can be examined by changing the measurement location.

PACS number: 43.70.Jt, 43.72.Ct

INTRODUCTION

Experimental measurement of the acoustic transfer function of the vocal tract is a crucial issue in understanding the speech production process, and it has been attempted by a number of researchers in the history of research on speech acoustics (Van den berg, 1955; Fant, 1962; Fujimura, et al, 1971). The fundamental problem is the mutual dependence of the source sound and the vocal tract transfer function in the acoustic domain. While one approach is to use features discovered for one of them to reveal properties of the other, in practice this is difficult to implement because both of them are inseparably coupled in the natural condition. The direct measurement of source sounds may be theoretically of less value if we consider the acoustic coupling effects on the source, and moreover this is difficult to achieve by non-invasive methods. For these reasons, direct measurement of the vocal tract transfer function is more practical and it has been attempted in a few limited situations. The first direct measurement of vocal tract transmission characteristics was reported by van den Berg (1955), who applied a sweeping pure-tone signal to the laryngeal fistula of a hemilaryngectomized subject. Fant (1962) applied a similar method to a normal subject by indirectly exciting his vocal tract through the skin above the larynx. The sound source he used was a small moving-iron type vibrator. By applying a sweeptone vibration directly on the external surface of the throat just above the larynx, he was able to record the frequency response curve of the vocal tract for different vowels and some nasalized vowels. Visual inspection of the continuous curves of frequency response near the formant peaks was used to estimate the bandwidths of the formants under a closed-glottis condition, and he derived a graph of the bandwidth values plotted over a frequency region up to about 4000 Hz. Fujimura & Lindqvist (1971) recorded vocal tract response by exciting the vocal tract with a sweeptone vibrator on the neck surface just above the larynx. They used an analysis-by-synthesis procedure to obtain reliable data for vowels, in particular of the formant bandwidths based on the recorded vocal tract responses. Castelli & Badin (1988) improved their method: the vocal tract was excited by a white noise signal instead of by a sweeptone signal in order to reduce measurement time. Djeradi, Guérin, Badin, and Perrier (1991) proposed the use of a known pseudo-random sequence instead of a white noise excitation. Their improvement further shortened the measurement time. All of these external excitation methods, however, have critical problems: it is evident that the recorded data of the vocal tract response are not independent from potential influence due to the transfer characteristics of the surface tissue and the cartilage of the larynx.

Taketa, Sato, and Nagata (1965) reported a method of measuring vocal tract transmission characteristics based on the principle of acoustical reciprocity: if volume velocity U_a at a point A produces sound pressure P_b at a point B, the same sound pressure will be induced at the point A if volume velocity U_a is placed at point B (c. f. Beranek, 1993). Using this principle they measured sound pressure in the vocal tract and volume velocity in front of the subject, and estimated transmission characteristics of the vocal tract for steady-state vowels. In this method the excitation sound source was placed in front of the subject. The subject was obliged to maintain a certain vocal tract shape with the vocal folds closed during the measurement.

Lindqvist-Gauffin and Sundberg (1976) conducted an experiment to measure transmission characteristics of the nasal tract, in which they used a probe tube sound source as the excitation source. In their study the sound source of a sweeptone signal was applied by a probe tube inserted into the nasopharynx, and the response of the nasal tract was recorded at the nostrils in a closed velopharyngeal port condition. The transfer function of the nasal tract, in particular pole-zeros caused by the paranasal sinuses, was examined by comparing the response of the nasal tract to the source signal. As the authors mentioned, a problem with the method is that extra zeros can be introduced depending on the position of the applied sound source. Since the outlet of the sound source probe was placed slightly away from the posterior wall of the nasopharynx, the portion from the probe's outlet to the posterior wall would act as an extra antiresonator to the measurement system.

The method proposed in this paper provides a direct measurement of the transmission characteristics of the vocal tract, in which the glottal sound can serve as the excitation sound source. The transfer function from an arbitrary location in the vocal tract to its radiation end (i.e., the lips and/or the nostrils) is estimated by measuring volume velocities at the location and at the radiation end. In this study, the volume velocities are measured using a pair of probe microphones for the internal vocal tract and by a regular microphone at the radiation end. Since this method in theory does not depend on the characteristics of the excitation sound source, the kinds and positions of the sound source have less influence on transfer function estimation. In addition to this, vocal tract transmission characteristics can be measured from short-period data selected from a natural utterance. The validity of the method was confirmed by conducting experiments on several acoustic tubes of known shapes. The accuracy of the method was tested by changing the experimental conditions.

I. THEORETICAL CONSIDERATIONS.

The vocal tract can be modeled as an acoustic tube of non-uniform, time-varying cross-sections. Since the wavelengths of the frequencies of our particular interest, i.e., for the range below 4000 Hz, are sufficiently long compared to the dimensions of the vocal tract, it is reasonable to assume a plane wave propagation along the axis of the acoustic tube (Rabiner, & Schafer, 1978). According to the conservation laws of mass, momentum, and energy, sound waves in the tube satisfy the following pair of equations:

$$-\frac{\partial p}{\partial x} = \rho \frac{\partial (u/A)}{\partial t} \tag{1}$$

$$-\frac{\partial u}{\partial x} = \frac{1}{\rho c^2}\frac{\partial (pA)}{\partial t} + \frac{\partial A}{\partial t} \tag{2}$$

where p is the sound pressure, u is the volume velocity, and A is the cross-sectional area of the vocal tract. Here, p, u, and A are functions of both the position x and time t. ρ is the air density and c is the sound velocity.

In the transmission line analogue, the vocal tract is computationally represented by a cascade concatenation of small sections. Each section of the vocal tract is approximately considered to be a uniform tube aligned perpendicular to the direction of wave propagation. A unit section of the vocal tract is specified by the cross-sectional area A and length l. Supposing the cross-sectional area A to be a constant over the time, the partial differential equations Eqs. (1) and (2) reduce to the form:

$$-\frac{\partial p}{\partial x} = \frac{\rho}{A}\frac{\partial u}{\partial t} \tag{3}$$

$$-\frac{\partial u}{\partial x} = \frac{A}{\rho c^2}\frac{\partial p}{\partial t} \tag{4}$$

The above equations indicate that sound pressure and volume velocity in any section can be derived from one another. Since the volume velocity is usually difficult to be measured directly, Equation (3) is more useful for obtaining the volume velocity. Equation (3) indicates that the time-derivative of the volume velocity is proportional to the sound pressure gradient in a section. If the sound pressure gradient at a position is approximated by the difference of sound pressures, the volume velocity at the position can be calculated using the sound pressures at two adjacent points near the position (c. f. Beranek, 1993). For realizing the computation, partial differential equation Eq. (3) is rewritten as a difference equation:

$$-\frac{p_2(n)-p_1(n)}{\Delta x}=\frac{\rho}{A}\frac{u(n)-u(n-1)}{\Delta t} \qquad (5)$$

where p_1 and p_2 are sound pressures measured simultaneously at two adjacent points separated by a distance Δx along the tract. The term u is the time sequence of the volume velocity, and it can be solved by Equation (6):

$$u(n)=-\frac{A}{\rho}\sum_{i=0}^{n}\frac{p_2(i)-p_1(i)}{\Delta x}\Delta t \qquad (6)$$

The volume velocity at the lips and/or the nostrils can be calculated using the radiation impedance and radiation sound pressure. According to the transmission line analogue, the radiation impedance is equivalent to the parallel connection of radiation resistance R_r and radiation inductance L_r (Rabiner, & Schafer, 1978). The volume velocity is solved with Eq. (7).

$$u_r(n)=\sum_{i=0}^{n}\left(\frac{p_r(i)}{L_r}+\frac{1}{R_r}\frac{p_r(i)-p_r(i-1)}{\Delta t}\right)\Delta t \qquad (7)$$
$$p(i)=0 \qquad (i<0)$$

where u_r and p_r are the volume velocity and sound pressure at the radiation end of the vocal tract. The transfer function T from the measurement location to the radiation end is represented by

$$T=\frac{U_r}{U} \qquad (8)$$

where U_r and U are the volume velocities at the radiation end and the measurement location, respectively.

Though we have an unknown quantity of cross-sectional area A in the above computations, as shown by Eq. (8) the relative magnitude of the volume velocities at input and output ends is the critical factor for the vocal tract transmission characteristics. The area A is only a constant coefficient in Eq. (6) for the computation of the volume velocity U, and does not affect the envelope of the frequency properties. The area A is, therefore, treated as unitary in these computations. However, this assumption has a limitation, that is, the distance Δx must be short enough so that the fundamental assumption is satisfied, namely, that the cross-sectional area is a constant within the distance Δx.

II. EXPERIMENTAL APPROACH

An experiment was conducted on acoustic tubes of known geometry to obtain their transmission characteristics using the proposed method. The transmission characteristics were then compared with theory-based computations in order to assess the validity of the method.

A. Setting for the experiment.

Four types of stiff plastic tubes, A, B, C, and D were used for the experiment. The geometries of the tubes are given in Table 1. All of the tubes had round uniform cross-sections with different branching configurations. A schematic diagram of the experimental setup is shown in Figure 1.

Table. 1

Figure 1

Two different types of excitation sound sources, a sweeptone signal in Setting (a) and the glottal sound source in Setting (b), were used in the experiment. In Setting (a), the sweeptone signal was generated by an FG-143 function generator to drive a SG-505FRP speaker, a horn driver unit. The acoustic tubes were connected to the throat of the horn driver unit for the measurement. Three microphones were used to measure the sound pressures inside and outside the tubes. Two B&K-4128 probe microphones, referred to as M1 and M2, were attached to each other to form parallel tubes, and were used for measuring the sound pressures within the tubes via rigid probe tubes of identical lengths. The other microphone was a B&K-4003, M3, used for obtaining the radiated sound pressure. The external microphone (M3) was placed about 3 cm away from the radiation end of the tube.

The glottal sound source is considered to be the most appropriate to measure the vocal tract transmission characteristics under natural conditions. In Setting (b), therefore, the glottal sound source was used as the excitation sound source to examine accuracy of this method. In the measurement, the subject put one end of the acoustic tube into his mouth and produced a steady-state vowel /o/. The sound pressures were recorded in three channels which were the same as those in Setting (a).

To obtain the sound pressure gradient, the positions of the two inlets of the probe tubes were separated by a certain distance along the direction of sound wave propagation. The distance between the two inlets is referred to as tip distance (TD). The location of measurement within a tube is given by the average value of the distances from each inlet of the probes to the radiation end of the acoustic tube. This location is called the "measurement location" (see Fig. 1). The portion from the measurement location to the radiation end of the acoustic tube is referred to as the "measurement portion." In this definition, the length of the measurement portion is equal to the value of the measurement location. The portion of the tube from the measurement location to the end connecting with the sound source is called the remaining portion.

B. Parameters used in the experiment.

The experiment was conducted in an anechoic room to take advantage of its low noise level. TD of the two inlets of the probe tubes of probe microphones was adjusted to a distance of 0.6 cm in both Settings (a) and (b). Two types of excitation sound sources were used: a sweeptone signal and the glottal sound source. The frequency range of the sweeptone signal was from 100 to 7500 Hz with sweep periods of 100 ms and 10 ms. These sweep rates were selected by considering the sharpness of the tube resonance for the sweep of 100 ms, and to test the validity of the method with a periodic sound source for 10 ms, which is close to the fundamental frequency of human speech sound. The sampling rate was 44.1 kHz, and the cut-off frequency was 7 kHz for this analysis. After the calculation of volume velocities, the cepstrum analysis was applied to low-pass filter the volume velocities obtained from the glottal sound and the sweeptone signal with a 10-msec sweep period. The window of the cepstrum analysis was 0.6 times as long as the period of the signals in order to eliminate periodicity effects.

C. Theoretical evaluations of resonance frequencies.

If the volume velocity obtained at the measurement location in the tube is treated as the input of the measurement portion, the transmission characteristics of the measurement portion are given by Eq. (8). For this treatment, the measurement portion can be considered to act as a piece of isolated tube independent from the remaining portion. This assumption can be confirmed by a comparison of the measured transmission characteristics of the measurement portion and theoretical values of a piece of isolated tube. The theoretical values that are the resonance frequencies were calculated from a uniform tube having the same geometry as the measurement portion, using Eq. (9).

$$f = \frac{(2n-1)c}{4(L+ar)} \quad (9)$$

where c is the sound velocity, n is the number of resonance peaks, L is the length of the measurement portion, a is the end correction coefficient, and r is the radius of the radiation end. The sound velocity was taken as 34,000 cm/s at a set temperature of 20 °C during the measurement. The correction for the radiation ends of the main tube and the branch was approximated by the product of the end correction coefficient and the radius. The end correction coefficient was assumed to be 0.8 for the radiation ends of both the main tube and the branch (Fant, 1970).

D. Calibrations

Some calibrations were made for the two probe microphones, M1 and M2, and for the volume velocities calculated with Eqs. (6) and (7), respectively. The results showed that the frequencies of zeros in transmission characteristics can be measured exactly even if the two microphones have a slight mismatch in frequency responses. However, the peaks in the frequency region below 1 kHz are sensitive to the mismatch between the two probe microphones. The properties of the two microphones were carefully calibrated in order to obtain accurate peaks over the low frequency region. Figure 2 shows the calibration coefficient for the microphones, which is the ratio of the amplitude of sound pressure recorded by microphone M1 to that recorded by microphone M2 when the same sound pressure signal was applied to both of them. In the calibration of volume velocities, the sound pressures were measured at the same measurement location in front of the radiation end. The result showed that the ratio of the volume velocities obtained from Eq (6) to those evaluated by Eq (7) resulted in a flat frequency spectrum within the region below 7 kHz when the measurement location is more than twice radius of the tube away from the radiation end.

Figure 2

III. RESULTS

Experiments were conducted on the acoustic tubes with the measurement portions ranging from 0 cm to 9 cm in a 1-cm increment for Tubes A, B, and C, and from 2 cm to 18 cm for Tube D, respectively. Tube A was measured under the conditions of both Settings (a) and (b), and the others were measured under Setting (a) only. Resonance and antiresonance properties were measured for all tubes.

A. Measurement of resonance properties

A uniform tube, Tube A, was used to test the accuracy of our method for the measurement of resonance properties. The transfer functions obtained from Tube A are shown in Fig. 3, where the excitation sound source was the sweeptone signal. The lengths of the measurement portions, i.e., the distances from the measurement location to the radiation end, are shown on the right side of the figure. The results show that the frequencies of the peaks of the transfer functions decrease regularly with increasing length of the measurement portion, and the number of peaks increases in the 0 to 7 kHz range. Some inaccurate results appear in the low frequency region below 100 Hz because the excitation source of sweeptone signal has no power in the region.

Figure 3

A consequent question addressed here is whether the measured transfer characteristics exactly represent those of the measurement portions or not. In other words, can each measurement portion be treated as a piece of isolated tube in this measurement system. To answer this question, a comparison was made between the measured results and theoretically computed values. The theoretical values were calculated using Eq. (9). The relative difference of the central frequency (RDF) of the peaks, i.e., the ratio of the difference between the measured and theoretical resonance frequencies to the theoretical value, was about 5% for the measurement portions whose lengths were shorter than or equal to 6 cm. The RDF was less than 3% for the measurement portions longer than 6 cm. The measurement portions show good agreement with the isolated tubes for the resonance frequencies, and the accuracy becomes higher with increasing length of the measurement portion. This is because of the relative decrease in the errors due to deviations from the ideal measurement location and the end correction. The results suggest that each measurement portion can be treated as a piece of isolated tube when the volume velocity at the measurement location is considered as the input.

For investigating the effect of diverse excitation sound sources on the measured results, the glottal sound of a male subject served as the excitation sound source in Setting (b). In this measurement, the subject produced a steady-state vowel /o/. The results obtained from the same uniform tube, Tube A, are shown in Figure 4. Compared with Fig. 3, no ripples appear in the low frequency region due to the glottal sound has enough power in that region. However, some dips appear near 4 kHz in the transfer function in Fig. 4. The dips are considered to be caused by some factors, which are either that the glottal source has insufficient power in the region near 4 kHz or that the energy was absorbed by some antiresonators in the vocal tract. They are discussed in Section IV in detail. Disregarding the region affected by the dips, the RDF of the peaks is about 5%. This RDF is almost the same as that obtained from the sweeptone sound source. The characteristics of excitation sound sources and the distances from the sound sources to the measurement position show no significant effect on the measurement results for the Settings (a) and (b), especially for the range below 4 kHz.

Figure 4

B. Measurement of antiresonance properties.

Vocal tract acoustic systems are of course more complex than a uniform tube; some have both resonance and antiresonance properties, such as vocal tract configurations for producing nasalization. Tubes B and C with different branches were used to test the validities of our method for an acoustical system which has both poles and zeros. Zeros of the tubes were expected to appear at the frequencies corresponding to the resonance frequencies of their branches. For Tube B, the branch was 3.4 cm long and 1.6 cm in diameter, and was located at 6.5 cm from the radiation end of the main tube. The antiresonance frequencies of this branch are theoretically 2104 and 6312 Hz in the region below 7 kHz, evaluated from Eq (9). Figure 5 illustrates the transfer functions obtained from Tube B, which were measured using the sweeptone signal of a 10-ms

Figure 5

sweep period. Two zeros, at about 2,154 Hz and 6,800 Hz, appear first in the transfer functions from the measurement location of 6 cm. The frequencies of the zeros correspond with the first and second antiresonance frequencies of the branch. The RDF between the measured value and the theoretical value of the zeros was about 2% for the first zero and 7% for the second zero. The appearance of the zeros coincides with the location of the branch. Peaks are regularly observed in the transmission characteristics obtained for the measurement locations from 2 to 5 cm, where there is no branch in the measurement portions. For these peaks, the RDF between the measured value and theoretical value was about 8% for the measurement portions.

Figure 6 illustrates transfer functions obtained from Tube C with a 6.0-cm-long branch. This branch had the same diameter and

Figure 6

location as those of Tube B. A series of zeros appear regularly in the transfer functions, and their antiresonance frequencies are 1,340, 4,053, and 6,699 Hz with an interval of about 2,680 Hz, twice the first antiresonance frequency. The RDF between the measured and theoretical values is about 2% for the three zeros. The branch of Tube C shows a larger effect on the transfer function than that of Tube B. The first zero at the measurement location of 5 cm splits the first resonance peak of the transfer function into two peaks. However, the effect of the zeros on the transfer function almost corresponds with the location of the branch. The resonance frequencies of the measurement portions shorter than 5 cm are consistent with the theoretical values within about 8% RDF, which is the same as that obtained from Tube B. The results obtained from Tubes B and C show that the measured transmission characteristics depend only on the geometry of the measurement portion, and they are not affected by the remaining portion from the measurement location to the sound source.

C. Measurement of an acoustic tube with more than one antiresonator.

When an acoustic system has more than one antiresonator such as the nasal cavity it is difficult for the traditional methods cited earlier to measure and locate the antiresonators, though some researchers (Fujimura & Lindqvist, 1971; Lindqvist-Gauffin & Sundberg, 1976) have attempted to measure them directly. To test our method's reliability for measuring and locating antiresonators in an acoustic system, an experiment was conducted on an acoustic tube, Tube D, which had two branches: one long and one short. The longer branch was 25.5 cm long, 1.2 cm in diameter, and located at 6.5 cm from the radiation end of the main tube. The shorter one was 3.4 cm long, 1.6 cm in diameter, and located at 12.5 cm from the radiation end. Flexible probe tubes were used for the measurement. The TD of the probe tubes was 0.6 cm. Figure 7 illustrates the transmission characteristics of the tube, which were obtained at measurement locations from 2 to 18 cm in a 2-cm increment. The

Figure 7

antiresonances caused by the longer branch, shown by vertical dashed lines, appear as regular zeros from the measurement location of 6 cm which corresponds to the location of the branch. The zeros caused by the shorter branch, indicated by arrows, appear first at the measurement location of 12 cm, the location of the branch. As shown in Fig. 7, the first antiresonance frequency of the shorter branch is close to the fourth antiresonance frequency of the longer branch, both observed separately. This result is consistent with a theoretical calculation that gives 2104 Hz for the first antiresonance frequency of the shorter branch and 2290 Hz for the fourth of the longer one. This suggests that even if the two branches have close antiresonance frequencies, they can be distinguished from one another by changing the measurement location. The RDFs of the zeros caused by the longer branch are smaller than 2% except for the first zero whose RDF is about 5%. The RDFs of the zeros are about 5%

for the shorter branch. An advantage of our method is shown in the above experiments: the antiresonance frequencies and the locations of the antiresonators of the measured acoustic system can be measured by changing the measurement location.

IV. PROBLEMS AND SOLUTIONS

The results obtained above showed the validity of our method for measuring the transmission characteristics, both peaks and zeros, of the acoustic tubes in the given conditions. Several instrumental problems, however, must be stressed for measuring vocal tract transmission characteristics using this method.

A. Effect of using bent probe tubes.

Different from the measurement of uniform tubes, the probe tubes of the microphones must be bent along the vocal tract to obtain the sound pressures within the vocal tract. A question here is how the bent probe tubes affect the accuracy of the measurement. To answer the question, we examined the effect in two conditions: one using straight probe tubes and the other using bent probe tubes. The probe tubes connected to Microphones M1 and M2 had an identical length of 23 cm and matched impedance to the microphones. The TD of the probe tubes was 0.6 cm. The bent probe tubes had an angle of 90 degrees at a position about 4 cm from the membrane of the microphones. Measurements were conducted at the measurement locations from 0 cm to 18 cm in a 1-cm interval. The measured results showed that the RDF was only 1% between the two conditions: straight and bent probe tubes. No significant effects were found on the measured results even if the probe tubes were bent at an angle of 90 degrees.

B. Effects of the TD.

As shown in Fig. 1, the inlets of the two probe microphones for obtaining sound pressure gradient were placed at two points with a distance Δx apart in an acoustic tube. It is known that the frequency resolution is correlated to the TD of the two microphones. In order to obtain an accurate volume velocity, the distance between the inlets of microphones should be adjusted according to the frequency region to be measured: a longer distance for a lower frequency region and a shorter distance for a higher frequency region (Beranek, 1993). Despite the advantage of a longer distance, useful TD length is constrained in the measurement of the vocal tract transfer function due to the dimensions of the vocal tract. The effect of the tip distance on the accuracy of the method is investigated by increasing the TD of the probe tubes from 0.3 to 2.0 cm. Flexible tubes with a 0.165-cm outer diameter and a 0.076-cm inner diameter were employed in this experiment. Measurements were obtained at the measurement locations from 0 cm to 19 cm in a 1-cm interval. Figure 8 illustrates the results obtained at the measurement location of 19 cm with different TDs of 0.3, 0.4, 0.5, 0.6, 1.5, and 2.0 cm. The results reveal almost the same transfer function for all of the TDs, and peaks with regular frequencies appear in the transmission

Figure 8

characteristics. Comparing the first resonance frequency shown in Fig. 8 with the theoretical values, the RDFs are smaller than 4% for TDs of 2.0, 1.5 and 0.6 cm and about 10% for the others. For the frequencies of high-order resonance peaks, the RDFs are less than 5% even for TDs shorter than 0.6 cm. A problem is that a shorter TD would cause deviations not only in the central frequency, but also in the amplitude of the resonance peaks when their frequencies are lower than 1 kHz. The bandwidth of the lower peaks shows a strong dependence on the TD. Therefore, a tradeoff exists between the TD and the measurement accuracy in the method. Observations of the zeros, however, indicated that the RDFs of the zeros hardly change when the TD of the probe tubes is reduced, while the peaks in the lower frequency range are sensitive to changes in the TD.

C. Some limitations in using the glottal sound source

As shown in Figure 4, local dips occurred in the transmission characteristics of the acoustic tube when the glottal sound served as the excitation source. To search for the causes of these dips, the volume velocities in the frequency region near the dips were investigated at both the measurement location and the radiation end. The data indicate that the volume velocity at the measurement location was close to the noise level in the frequency region. There are two possibilities about the phenomenon: one is that the glottal sound has insufficient power in the region; the other is that sound energy may be absorbed by some antiresonators in the vocal tract. Considering the latter one, a possible antiresonator for the vocal tract configuration producing vowel /o/ may be the pyriform fossa, which consists of two airy pouches lying on each side of the larynx as a part of the hypopharyngeal cavity. The shape of the pyriform fossa is goblet-like with slightly constricted openings obliquely bordered by the aryepiglottis folds. Acoustically, the pyriform fossa can serve as two respective side-branches to the main vocal tract. With this assumption, we analyzed volumetric MRI images to obtain the geometry of the fossa for the subject. The antiresonance frequencies of the two side-branches, which were obtained using Eq. (9), were about 3,810 and 3,950 Hz respectively. These frequencies were almost consistent with those of the dips. When the same investigation was implemented for some other subjects, the antiresonance frequencies obtained from the geometry of the pyriform fossa were also consistent with the dips near 4 kHz in spectrum of their speech sound. This finding leads us to believe that the pyriform fossa can be an important cause of the dips. The result implies a limitation of the glottal sound used as an excitation sound source. It seems that measurement of transmission characteristics of an acoustic tube becomes less accurate in the frequency regions where the glottal source is low in energy.

V. CONCLUSIONS

A new method for direct measurements of transmission characteristics of the vocal tract has been proposed, and has been experimentally studied using acoustical tubes of known geometry. The method yielded an accuracy of about 4% RDF for the peaks and 2% for the zeros of the acoustic tubes, obtained by comparing the measured transmission characteristics to the theoretical computation. Compared with previous methods the present method has two major advantages: first, the characteristics and positions of the excitation sound source have less influence on the measurement of the frequencies of the peaks and zeros; second, the measured results obtained from the measurement portion exactly represent acoustic properties of the measurement portion from the measurement location to the radiation end, and are not affected by the remaining portion. By taking advantage of our method, the glottal sound can serve as the excitation sound source for the direct measurement. The spectrum details of the vocal tract transmission characteristics can be examined by changing the measurement location. For example, the antiresonance frequencies and the locations of the paranasal cavities can be measured and located using this method.

Our method, in principle, estimates vocal tract transmission characteristics using the volume velocities inside the vocal tract and at the radiation end. The measurement accuracy of this method to a great extent depends on the accuracy of the volume velocity inside the vocal tract. In order to obtain higher accuracy of the volume velocity, especially in the lower frequency region, a longer TD would be required. In application to a real vocal tract, however, the TD should be shorter than 0.8 cm. Considering the tradeoff between measurement accuracy and the TD, a TD length between 0.4 and 0.6

cm is recommended for measuring vocal tract transmission characteristics. Assuming that the TD is 0.6 cm, for example, the accuracy can be expected to be about 4% in the RDF for the peaks and about 2% for the zeros for nasalized and non-nasalized vowels.

While this method can be used to measure vocal tract transmission characteristics using natural utterances, a few disadvantages have been pointed out with respect to the use of the glottal sound source. They include the harmonic structure of the fundamental frequency and lack of sufficient energy in some frequency regions. The harmonic structure may cause deterioration of some spectral details of vocal tract transmission characteristics, and the lack of sufficient energy may result in inaccurate measurement. Our method showed an advantage for examining local properties of the vocal tract. Measuring an entire transfer function of the vocal tract quires the probe to be inserted as near the vocal folds as possible. An instrumental problem with such a measurement is that a surface anesthesia would be necessitated for the larynx. As shown in these experiments, our method can be usefully applied in measuring vocal tract transmission characteristics, especially for investigating the distribution of the pole-zeroes of nasalization.

ACKNOWLEDGMENTS

The authors would like to thank Hiroyuki Hirai for his helpful discussions and comments. We are also grateful to Mark Tiede for his help in this translation.

Beranek, L. L. (1993). *Acoustical measurements*, AIP, Massachusetts, p.261 (Revised edition).

Djeradi, A., Guérin, B., Badin, P., and Perrier, P. (1991). "Measurement of the acoustic transfer function of the vocal tract: a fast and accurate method," Journal of Phonetics, 19,387-395.

Fant, G. (1962). "Formant bandwidth data," STL-QPSR, 1, 1-2.

Fant, G. (1970). *Acoustic theory of speech production*, Mouton & Co., Netherlands, p.36 (Second edition).

Fujimura, O., and Lindqvist J. (1971). "Sweep-tone measurements of vocal tract characteristics," J. Acoust. Soc. Ame., 49, 2, 541-558.

Lindqvist-Guaffin, J., and Sundberg J. (1976). "Acoustic properties of the nasal tract," Phonetica, 33, 161-168.

Rabiner, L. R., and Schafer, R. W. (1978). *Digital processing of speech signals,* Prentice-Hall, Inc., New Jersey.

Takeda, T., Sato, Y. and Nagata K. (1965). " Measurement of vocal tract transfer function," 5th Congress International D'Acoustique, Liege.

Van den Berg, J. (1955). "Transmission of the vocal cavities," J. Acoust. Soc. Ame., 27, 161-168.

Table caption

Table 1 Geometries of the tubes used in the experiment.

Figure captions

Fig. 1 Schematic diagram of set-up for measuring the transfer function of acoustic tubes.

(a) Excitation sound source is a sweeptone signal.

(b) Excitation sound source is the glottal sound source.

Fig. 2 Calibration coefficient between two probe microphones.

Fig. 3 Transfer functions obtained from a uniform tube using a sweeptone sound source. (The measurement locations from the radiation end are shown on the right side.)

Fig. 4 Transfer functions obtained from a uniform tube using the glottal sound source. (The measurement locations from the radiation end are shown on the right side.)

Fig. 5 Transfer functions obtained from Tube B with a 3.4-cm-long branch using a sweeptone sound source. (The measurement locations from the radiation end are shown on the right side. Zeros are indicated by dashed lines)

Fig. 6 Transfer functions obtained from Tube C with a 6-cm-long branch using a sweeptone sound source. (The measurement locations from the radiation end are shown on the right side. Zeros are shown by dashed lines.)

Fig. 7 Transfer functions obtained from a test with two branches. (The measurement locations from the radiation end are shown on the right side. Dashed lines indicate the zeros caused by a 25.5-cm-long branch. The arrows show the zeros caused by a 3.4-cm-long branch.)

Fig. 8 Transfer functions obtained at the measured distance of 19 cm using different tip displacements. (The tip distances of probe tubes of microphones are shown on the right side.)

Table 1 Geometries of the tubes used for the experiment. (Unit:cm)

Tubes	Main Tubes		Branch 1			Branch 2		
	Leng.	Diam.	Loca.	Leng.	Diam.	Loca.	Leng.	Diam.
Tube A	21.0	2.0	-	-	-	-	-	-
Tube B	11.7	2.2	6.5	3.4	1.6	-	-	-
Tube C	11.7	2.2	6.5	6.0	1.6	-	-	-
Tube D	19.0	2.2	6.5	25.5	1.2	12.5	3.4	1.6

*Leng.:Length; Diam.: Diameter;
Loca.: Location of branches from the radiation end of the main tube.

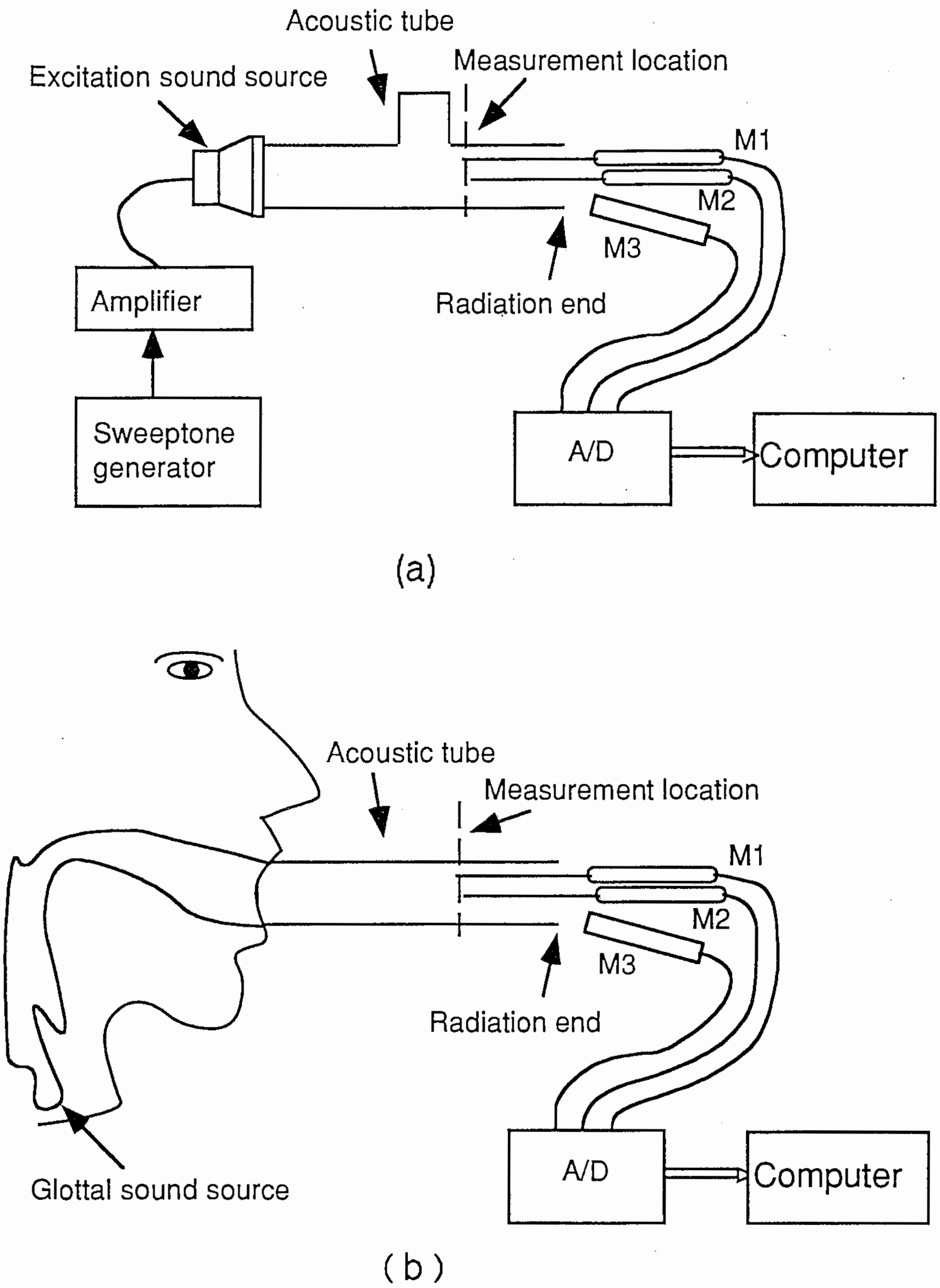

Fig. 1 Schematic diagram of set-up for measuring the transfer function of acoustic tubes.
(a) Excitation sound source is a sweeptone signal.
(b) Excitation sound source is the glottal sound source.

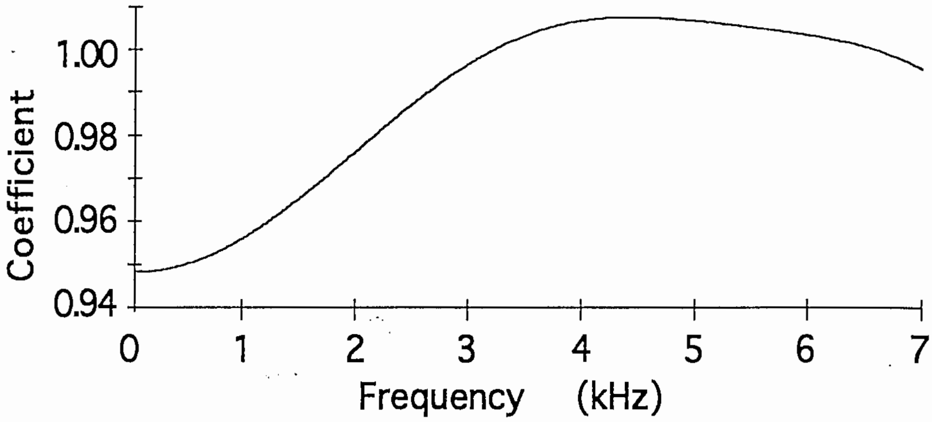

Fig. 2 Calibration coefficient between two probe microphones.

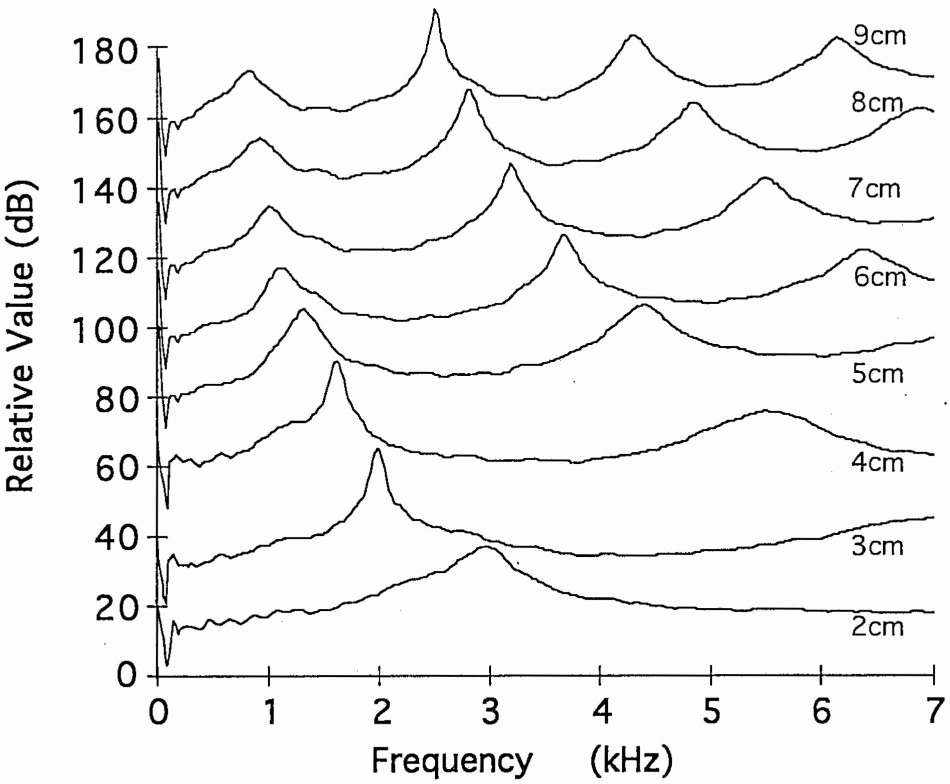

Fig. 3 Transfer functions obtained from a uniform tube using a sweeptone sound source. (The measurement locations from the radiation end are shown on the right side.)

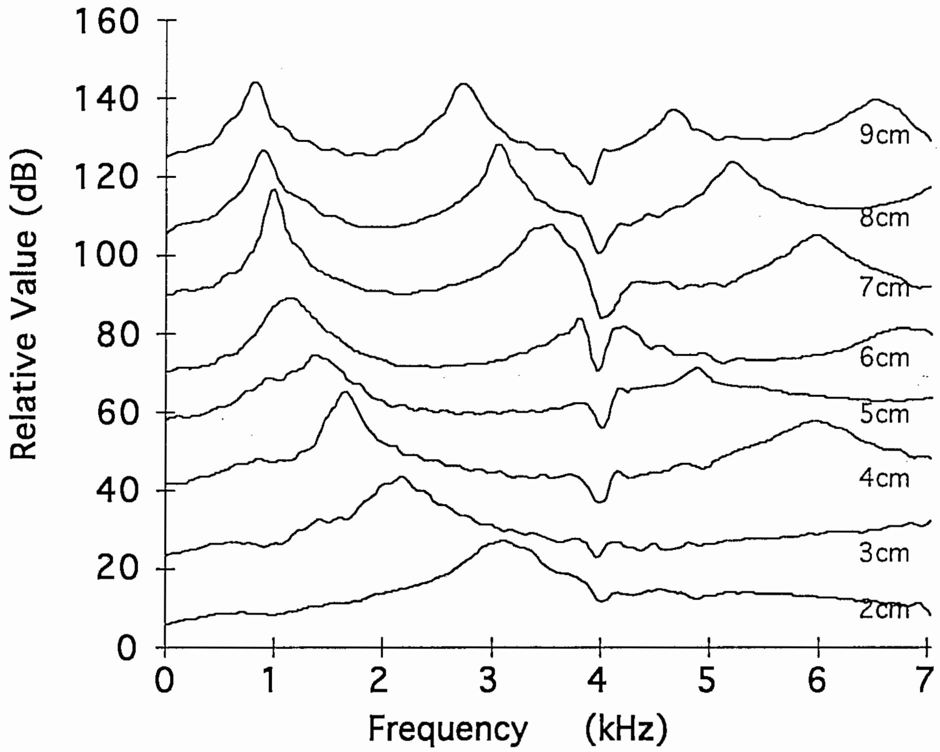

Fig. 4 Transfer functions obtained from a uniform tube using the glottal sound source. (The measurement locations from the radiation end are shown on the right side.)

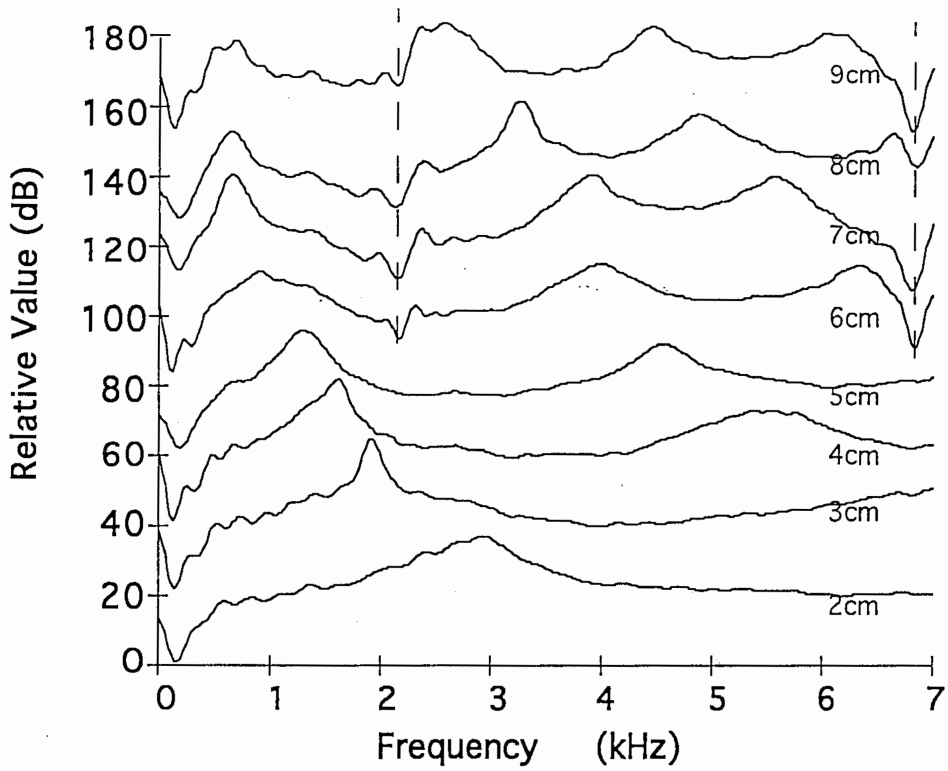

Fig. 5 Transfer functions obtained from Tube B with a 3.4-cm-long branch using a sweeptone sound source. (The measurement locations from the radiation end are shown on the right side. Zeros are indicated by dashed lines)

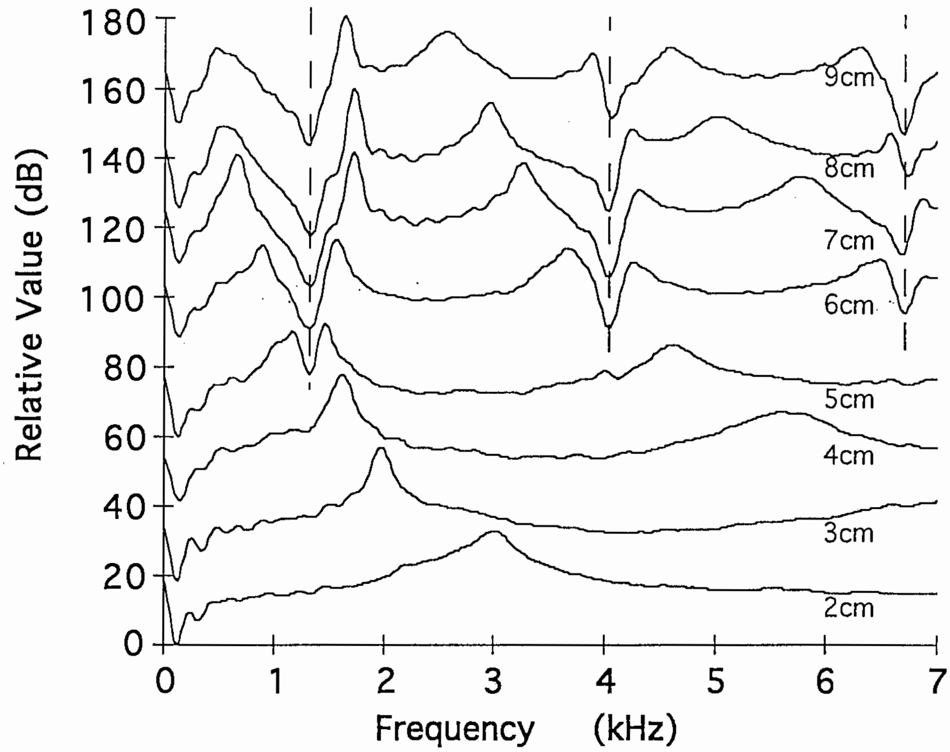

Fig. 6 Transfer functions obtained from Tube C with a 6-cm-long branch using a sweeptone sound source.
(The measurement locations from the radiation end are shown on the right side. Zeros are shown by dashed lines.)

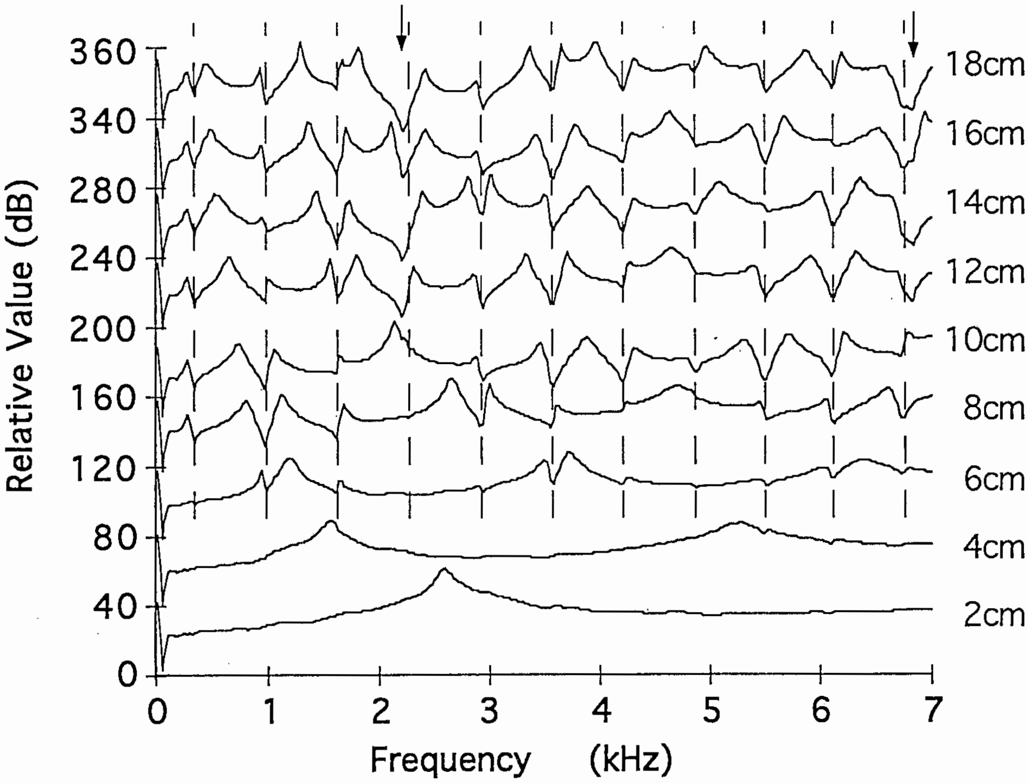

Fig. 7 Transfer functions obtained from a test tube with two branches. (The measurement locations from the radiation end are shown on the right side. Dashed lines indicate the zeros caused by a 25.5-cm-long branch. The arrows show the zeros caused by a 3.4-cm-long branch.)

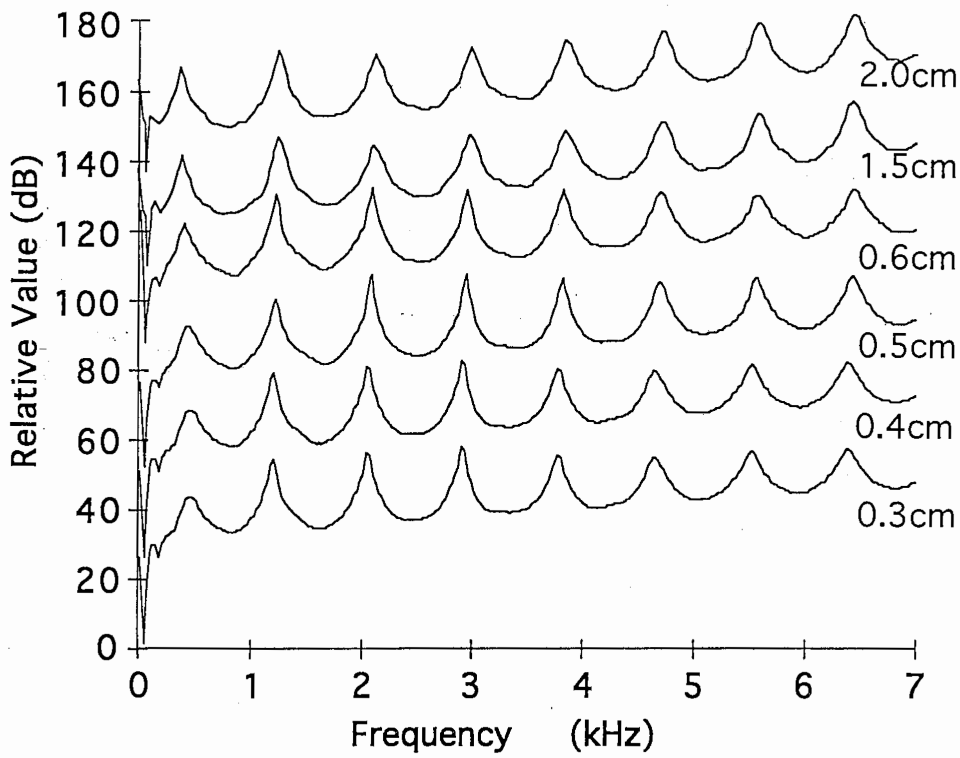

Fig. 8 Transfer functions obtained at the measured distance of 19cm using different tip displacements. (The tip distances of probe tubes of microphones are shown on the right side.)