A parametric method of computing acoustic characteristics of simplified three-dimensional vocal-tract model with wall impedance

Kunitoshi Motoki

Department of Electronics and Information Engineering, Hokkai-Gakuen University, S–26 W–11, Chuo-ku, Sapporo, 064–0926 Japan

(Received 28 May 2012, Accepted for publication 1 November 2012)

Abstract: A method of computing the acoustic characteristics of a simplified three-dimensional vocal-tract model with wall impedance is presented. The acoustic field is represented in terms of both plane waves and higher order modes in tubes. This model is constructed using an asymmetrically connected structure of rectangular acoustic tubes, and can parametrically represent acoustic characteristics at higher frequencies where the assumption of plane wave propagation does not hold. The propagation constants of the higher order modes are calculated taking account of wall impedance. The resonance characteristics of the vocal-tract model are evaluated using the radiated acoustic power. Computational results show an increase in bandwidth and a small upward shift of peaks, particularly at lower frequencies, as already suggested by the one-dimensional model. It is also shown that the sharp peaks at higher frequencies are less sensitive to the values of wall impedance even though the attenuation of the higher order modes is larger than that of plane waves.

Keywords: Vocal-tract model, Higher order modes, Mode matching technique, Wall impedance

PACS number: 43.70.Bk [doi:10.1250/ast.34.113]

1. INTRODUCTION

The acoustic analysis of three-dimensional vocal-tract models at higher frequencies has been performed by many researchers using the geometrical data obtained by magnetic resonance imaging (MRI). A finite-element method (FEM) is a standard numerical simulation technique for investigating the effects of the fine three-dimensional structure of the vocal tract, such as small dips or branches [1–3]. However, the FEM requires a large amount of time not only for the computation itself, but also for the creation of finite-element meshes. A finite-difference time-domain method (FDTD) is also used for the analysis of the three-dimensional vocal tract [4]. As the computation nodes can be specified at the same position as the voxels of the MRI data, the FDTD has an advantage for the creation of the entire configuration of the vocal tract. The FDTD is also suitable for parallel computing, but encounters difficulties in representing the impedance properties of the boundaries.

On the other hand, a parametric method to compute the acoustic characteristics at higher frequencies, where the assumption of plane wave propagation does not hold, has been developed by using higher order modes [5–8]. The vocal tract is approximated by a cascaded structure of rectangular acoustic tubes. The acoustic field in this model is represented in terms of both plane waves and higher order modes. Although the performance in representing the detailed structure of the vocal tract is limited, this model can represent the effects of the transverse dimension of the vocal tract at higher frequencies, and has the advantage of fast computation since this method can be regarded as an extension of the well-known one-dimensional vocal-tract model.

In this study, wall impedance is introduced in the computation of the propagation constant of the higher order modes. Using the power radiated in a free space, the transfer characteristics of the proposed model are evaluated for different values of wall impedance. Computational results show an increase in the bandwidth and a small upward shift of peaks, particularly at lower frequencies, as already suggested by the one-dimensional model. It is also shown that the sharp peaks at higher frequencies are less sensitive to the values of wall impedance even though the attenuation of the higher order modes is larger than that of plane waves.

2. SIMPLIFIED THREE-DIMENSIONAL MODEL

2.1. Overview of the Model

A cascaded structure of rectangular acoustic tubes,
connected asymmetrically with respect to their axes, as illustrated in Fig. 1, is introduced as a simplified approximation of the vocal-tract geometry. The width, height, length, and the position of the axis of each tube can be specified independently. A sound source can be specified as an arbitrary vibrating surface at the entrance of the first section. The last section, corresponding to the mouth, is open to free space. The radiation of sound is taken into account. The wall of each rectangular tube has yielding properties that can be specified by the wall impedance.

2.2. Propagation Constant of Higher Order Modes

A sound pressure \( p(x,y,z) \), \( z \) being the direction of the tube axis, and a particle velocity \( \psi(x,y,z) \) can be expressed as

\[
p(x,y,z) = j\omega\rho\phi(x,y,z),
\]

(1)

\[
\psi(x,y,z) = -\nabla\phi(x,y,z),
\]

(2)

where \( \omega \) and \( \rho \) are angular frequency and air density, respectively. \( \phi(x,y,z) \) is a velocity potential satisfying the Helmholtz equation.

\[
\nabla^2\phi(x,y,z) + k^2\phi(x,y,z) = 0
\]

(3)

where \( k = \omega/c \), \( c \) being sound speed, is the wave number. The three-dimensional acoustic field in each tube can be represented as an infinite series of higher order modes as

\[
\phi(x,y,z) = \sum_{m,n=0}^{\infty} \left[ a_{mn}\exp(-\gamma_{mn}z) + b_{mn}\exp(\gamma_{mn}z) \right] \psi_{mn}(x,y),
\]

(4)

where \( m \) and \( n \) stand for the indices of the higher order modes in the \( x \)- and \( y \)-directions, respectively, and \( a_{mn} \) and \( b_{mn} \) are constants determined from the geometry of adjacent tubes. \( \gamma_{mn} \) and \( \psi_{mn}(x,y) \) are the propagation constant and normal function (eigenfunction), respectively. \( \psi_{mn}(x,y) \) can be written as

\[
\psi_{mn}(x,y) = H_{cs}(m,\gamma_{cs,m})H_{cy}(n,\gamma_{cy,n}),
\]

(5)

where \( \gamma_{cs,m} \) and \( \gamma_{cy,n} \) are propagation constants in the \( x \)- and \( y \)-directions, respectively, and \( H_{cs}(q,\theta) \) is defined as

\[
H_{cs}(q,\theta) = \begin{cases} 
\mu_q \cosh(\theta) & q = 0, 2, 4, \ldots \\
\mu_q \sinh(\theta) & q = 1, 3, 5, \ldots 
\end{cases}
\]

(6)

where \( \mu_q = 1 \) (\( q = 0 \)) and \( \mu_q = \sqrt{2} \) (\( q \geq 1 \)). \( \gamma_{cs,m} \), \( \gamma_{cy,n} \), and \( \gamma_{mn} \) are required to satisfy

\[
\gamma_{cs,m}^2 + \gamma_{cy,n}^2 + \gamma_{mn}^2 + k^2 = 0.
\]

(7)

\( \gamma_{mn} \) is determined by specifying the boundary condition on the wall, and has an attenuation constant due to the wall impedance. The wall impedance \( Z_w \) is defined as the ratio of the sound pressure \( p \) to the particle velocity component \( v_n \) normal to the wall surface. By applying the theory of lined ducts [9,10], the wall boundary condition can be represented separately for even and odd modes in the \( x \)-direction.

Even modes:

\[
\left. \frac{\partial p}{\partial x} \right|_{x=0} = 0,
\]

(8)

\[
Z_w = \left. \frac{p}{\pm v_n} \right|_{x=\pm L_x/2} = -j\omega\rho \coth \left( \gamma_{cs,m} \frac{L_x}{2} \right).
\]

(9)

Odd modes:

\[
\left. p \right|_{x=0} = 0,
\]

(10)

\[
Z_w = \left. \frac{p}{\pm v_n} \right|_{x=\pm L_x/2} = -j\omega\rho \tanh \left( \gamma_{cs,m} \frac{L_x}{2} \right).
\]

(11)

where \( L_x \) is the sectional dimension of the rectangular tube in the \( x \)-direction, as illustrated in Fig. 2. From Eqs. (8) and (9), \( \gamma_{cs,m} \) is obtained as

\[
\gamma_{cs,m} = \begin{cases} 
\frac{3}{\pi^2} \sqrt{\frac{2k\bar{Y}}{L_x}} & m = 0 \\
\frac{j\pi}{L_x} \left( m + \frac{2kL_x\bar{Y}}{mn^2} \right) & m = 1, 2, 3, \ldots
\end{cases}
\]

where \( \bar{Y} \) is the normalized wall admittance defined by \( \bar{Y} = \rho c/Z_w \), and \( |\bar{Y}| \ll 1 \) is assumed. Similar relations can be written in the \( y \)-direction:

\[
\gamma_{cy,n} = \begin{cases} 
\frac{\sqrt{2}}{\pi^2} \sqrt{\frac{2k\bar{Y}}{L_y}} & n = 0 \\
\frac{j\pi}{L_y} \left( n + \frac{2kL_y\bar{Y}}{mn^2} \right) & n = 1, 2, 3, \ldots
\end{cases}
\]
Then \( \gamma_{z,\text{mn}} \), the propagation constant of mode \((m, n)\) in the \(z\)-direction, is calculated as
\[
\gamma_{z,\text{mn}} = \alpha_{z,\text{mn}} + j\beta_{z,\text{mn}} = \sqrt{-(\gamma_{x,\text{mn}}^2 + \gamma_{y,\text{mn}}^2 + k^2)}.
\] (12)
\( \alpha_{z,\text{mn}} \) and \( \beta_{z,\text{mn}} \) are attenuation and phase constants, respectively. The mode \((0, 0)\) represents a plane wave; others are higher order modes.

The rigid wall condition corresponds to \(|\tilde{Y}| = 0\), and \( \gamma_{z,\text{mn}} \) is reduced to
\[
\gamma_{z,\text{mn}} = \sqrt{\left(\frac{m\pi}{L_x}\right)^2 + \left(\frac{n\pi}{L_y}\right)^2} - k^2.
\] (13)
Under this condition, the cutoff frequency \( f_{c,\text{mn}} \) of mode \((m, n)\) is calculated as the frequency where \( \gamma_{z,\text{mn}} = 0 \).
\[
f_{c,\text{mn}} = \frac{c}{2\pi} \sqrt{\left(\frac{m\pi}{L_x}\right)^2 + \left(\frac{n\pi}{L_y}\right)^2}.
\] (14)
The higher order modes are propagative above \( f_{c,\text{mn}} \) and evanescent, \( \beta_{z,\text{mn}} = 0 \), below \( f_{c,\text{mn}} \). The evanescent modes are mainly excited at regions of discontinuity in the shape, and are localized in the vicinity of the excitation regions. In order to represent the variation of the acoustic field in the vicinity of regions with a rapid change in sectional shape, several higher order modes, not only propagative modes but also evanescent modes, must be considered. In the case of a soft wall condition, \( \tilde{Y} \neq 0 \), the exact cutoff frequency cannot be obtained as \( \gamma_{z,\text{mn}} \) does not become 0. However, as \( \tilde{Y} \) is small, it is assumed that the resonances due to the higher order modes will appear at frequencies above \( f_{c,\text{mn}} \). If no higher order mode is considered, i.e., there are only plane waves, this model is identical to the ordinary one-dimensional model of speech production.

### 2.3. Mode Expansion and Coupling

From Eqs. (1), (2), and (4) the sound pressure \( p(x, y, z) \) and the \( z\)-direction particle velocity \( v_z(x, y, z) \) in each tube are expressed as follows:
\[
p(x, y, z) = j\omega \rho \sum_{m=0}^{\infty} \left( a_{mn} \exp(-\gamma_{z,\text{mn}}z) + b_{mn} \exp(\gamma_{z,\text{mn}}z) \right) \psi_{nm}(x, y),
\]
\[
\approx \Psi^T(x, y) \left[ D(-z) a + D(z) b \right],
\] (15)
\[
v_z(x, y, z) = -\frac{\partial \phi(x, y, z)}{\partial z}
\approx \tilde{\psi}^T(x, y) Z_C^{-1} \left[ D(-z) a - D(z) b \right].
\] (16)
In the matrix notation in Eqs. (15) and (16), the infinite series are truncated to a certain value. \( a, b, \) and \( \psi(x, y) \) are column vectors composed of \( j\omega \rho a_{mn}, j\omega \rho b_{mn}, \) and \( \psi_{mn}(x, y), \) respectively. \( a \) and \( b \) are determined by the boundary conditions at the two ends of each tube. \( D(z) \) and \( Z_C \) are defined as
\[
D(z) = \text{diag} \left[ \exp(\gamma_{z,\text{mn}}z) \right],
\]
\[
Z_C = \text{diag} \left[ Z_{C,\text{mn}} \right],
\] (17)
where \( Z_{C,\text{mn}} \) is the characteristic impedance of mode \((m, n)\) and is defined as
\[
Z_{C,\text{mn}} = \frac{j\omega \rho}{\gamma_{z,\text{mn}}}
\] (18)
From Eqs. (15) and (16), a modal sound-pressure vector \( P \) and a modal particle velocity vector \( V \) can be defined as
\[
P = D(-z) a + D(z) b,
\] (19)
\[
V = Z_C^{-1} \left[ D(-z) a - D(z) b \right],
\] (20)
and are functions of the coordinate \( z \) only. If each component of \( P \) and \( V \) is regarded as a voltage and a current at position \( z \), each higher order mode can be represented by an equivalent electrical transmission line.

Hereafter, the subscript \( i \) is used to represent the variables in section \( i \). By using a mode matching technique, the mode coupling at the junction between sections \( i \) and \( i+1 \) can be expressed as [11]
\[
P_i = \Psi_{i,i+1} P_{i+1},
\] (21)
\[
\Psi_{i,i+1}^T V_i = V_{i+1},
\] (22)
where the coupling matrix \( \Psi_{i,i+1} \) is calculated as
\[
\Psi_{i,i+1} = \frac{1}{S_i} \int_{S_i} \Psi(x, y) \Psi^T_{i+1}(x, y) dS.
\] (23)
\( S_i \) is the area of section \( i \). \( S_i \) is assumed to be smaller than that of section \( i+1 \). If the smaller area is for section \( i+1 \), all suffixes \( i \) and \( i+1 \) should be exchanged. Equations (21) and (22) indicate that the coupling coefficient matrix \( \Psi_{i,i+1} \) can be simply regarded as a matrix representing the transformation ratio of a multiport ideal transformer in an equivalent electrical circuit. As one higher order mode corresponds to one transmission line, the equivalent electrical circuit of a tube and junction illustrated in Fig. 3 can be obtained.

### 2.4. Impedance Transformation

Acoustic characteristics of the vocal-tract model can be calculated from input impedance matrices at each section. Starting with a given load impedance matrix, which is a radiation impedance matrix, an impedance transformation is necessary to obtain the input impedance matrices for all sections. Kergomard [11] presented a propagating modes method for the impedance transformation where all evanescent modes are terminated at the junction. Pagneux et al. [12] discussed impedance transformation techniques for junctions that can even be carried to the limit of an infinitesimal step junction. In this section, the formulation of an impedance transformation suitable for the simplified three-dimensional vocal-tract model is
presented. It should be noted that the evanescent higher order modes do not propagate. However, they influence the resonance characteristics through the mode coupling between plane waves and the evanescent modes at the junction. Moreover, if two junctions are located very close together, the evanescent modes can be related to a power transmission. Thus, the evanescent modes are considered as either terminated (lumped element) at the junction or connected (line element) in the tube.

In the equivalent electrical circuit in Fig. 3, the terminated lumped elements are for the evanescent modes of very high order, shown as elements filled in black connected to the ground, and are located on both sides of the ideal transformer \( \Psi_{i+1,i} \). These elements are the same as the characteristic impedances in Eq. (18) and have no direct coupling to the adjacent ideal transformer. The evanescent modes with lower cutoff frequencies are shown as line elements filled in black connecting two ideal transformers. Depending on the distance between the junctions, possible interactions of the evanescent modes are represented by these line elements.

With the above treatment, the number of modes at each section can be specified independently. In particular, for narrow sections, only plane waves can be considered as “related to the transmission,” whereas for wide sections, several modes can be included.

A radiation impedance matrix is first used to start the impedance transformation. The generalized modal radiation impedance matrix \( Z_{\text{rad}} \), which relates the modal sound-pressure vector \( P \) and particle velocity vector \( V \) on the radiating surface of a rectangular opening with the higher order modes, is given as follows [13]:

\[
Z_{\text{rad}} = [Z_{mn,pq}],
\]

\[
P = Z_{\text{rad}} V,
\]

\[
Z_{mn,pq} = \frac{j k p c}{2 \pi S} \int \int_S \psi_m(x,y) \psi_n(x',y') e^{-jkr} dS \ ds,
\]

where \( S \) is the area of the last section corresponding to the mouth. \( Z_{00,00} \) is the radiation impedance used in the plane wave theory.

Input impedances looking toward the load at the right and left ends of the section \( i \) are defined as

\[
P_i^{(R)} = Z_i^{(R)} V_i^{(R)} \quad \text{and} \quad P_i^{(L)} = Z_i^{(L)} V_i^{(L)},
\]

where superscripts \( (R) \) and \( (L) \) are used to denote the quantities at the right (lip side) and left (glottis side) ends of the tube. From Eqs. (19) and (20), the impedance transformation through a tube (from \( Z_i^{(R)} \) to \( Z_i^{(L)} \)) is obtained as

\[
Z_i^{(L)} = (D_{C,i} Z_i^{(R)} + D_{S,i} Z_{C,i}) (D_{S,i} Z_i^{(R)} + D_{C,i} Z_{C,i})^{-1} Z_{C,i},
\]

where \( D_{C,i} = [D_i(L_{c,i}) + D_i(-L_{c,i})]/2 \) and \( D_{S,i} = [D_i(L_{c,i}) - D_i(-L_{c,i})]/2, L_{c,i} \) being the length of section \( i \).

The impedance transformation through a junction (from \( Z_{i+1}^{(L)} \) to \( Z_i^{(R)} \)) is given from the well-known circuit theory as

\[
Z_i^{(R)} = \Psi_{i,i+1} Z_{i+1}^{(L)} \Psi_{i,i+1}^T.
\]

The size of the coupling coefficient matrix \( \Psi_{i,i+1} \) is dependent on the number of modes considered, and is not always square. The number of modes selected for the calculation of \( \Psi_{i,i+1} \) can be large. It may be, however, limited to 5 or 6 at most. Some of these modes are considered for transmission, while the others are terminated with their characteristic impedances, as illustrated in Fig. 3. Some input impedance matrices are also defined in Fig. 3 as follows.

\( Z_i^{(R)} \): Input impedance matrix looking toward the load side at the right end for the modes considered for transmission in the \( i \)th section.

\( Z_i^{(L)} \): Input impedance matrix looking toward the load side at the left end for the modes considered for transmission in the \( i \)th section.

\( Z_i^{(R)} \): Diagonal impedance matrix to represent the terminated port at the right end of the \( i \)th section.

\( Z_i^{(L)} \): Diagonal impedance matrix to represent the terminated port at the left end of the \((i+1)\)th section.

\( Z_i^{(R)} \): Input impedance matrix looking toward the load side at the right end for all modes in the \( i \)th section.

\( Z_i^{(L)} \): Input impedance matrix looking toward the load side at the left end for all modes in the \((i+1)\)th section.

\[ Z_{mn,pq} = \frac{j k p c}{2 \pi S} \int \int_S \psi_m(x,y) \psi_n(x',y') e^{-jkr} dS \ ds,
\]

\[
r = \sqrt{(x-x')^2 + (y-y')^2},
\]
The problem to solve for the impedance transformation at each section is to express $Z_{i+1}^{(L)}$ in terms of $Z_i^{(R)}$, $Z_{i+1}^{(R)}$, and $Z_i^{(L)}$. $Z_i^{(R)}$ and $Z_{i+1}^{(R)}$ are simply diagonal matrices composed of characteristic impedances used for the termination of the ideal transformer. $Z_{i+1}^{(L)}$ is written as

$$Z_{i+1}^{(L)} = \begin{bmatrix} Z_{i+1}^{(L)} & 0 \\ 0 & Z_i^{(L)} \end{bmatrix}. \quad (29)$$

Then from Eq. (28), expanding the size of $\psi_{i+1}$ by adding some modes so that $\psi_{i+1}$ becomes square, the following relation holds.

$$\tilde{Z}_i^{(R)} = \psi_{i+1}Z_{i+1}^{(L)}\psi_i^{T} \quad (30)$$

$\tilde{Z}_i^{(R)}$ can be decomposed into submatrices as

$$\tilde{Z}_i^{(R)} = \begin{bmatrix} \tilde{Z}_{i,11}^{(R)} & \tilde{Z}_{i,12}^{(R)} \\ \tilde{Z}_{i,21}^{(R)} & \tilde{Z}_{i,22}^{(R)} \end{bmatrix} \quad (31)$$

where the size of $\tilde{Z}_{i,11}^{(R)}$ corresponds to the number of modes considered for transmission in the $i$th section. Then $Z_i^{(R)}$ is calculated as

$$Z_i^{(R)} = \tilde{Z}_{i,11}^{(R)} - \tilde{Z}_{i,12}^{(R)}(\tilde{Z}_{i,22}^{(R)} + \tilde{Z}_i^{(R)})^{-1}\tilde{Z}_{i,21}^{(R)}. \quad (32)$$

Finally, $Z_i^{(L)}$ is obtained by substituting $Z_i^{(R)}$ into Eq. (27). Note that $Z_i^{(L)}$ and $Z_{i+1}^{(L)}$ can be different in size. For a very narrow section, it may be wise to consider only plane waves to ensure the stability of the actual computation. This does not mean that higher order modes are not considered throughout the entire configuration. The impedance for plane waves at the end of the narrow section can be influenced by the higher order modes in the next large section.

Repeating the above procedure section by section, the given radiation impedance matrix $Z_{rad}$ in Eq. (25) is transformed into the input impedance matrix $Z_1^{(L)}$ of the first section. At the entrance of each section, a vibrating surface is given as a velocity distribution $v_g(x, y)$. The modal particle velocity vector at the left end of the first section, $V_1^{(L)}$, can be obtained as

$$V_1^{(L)} = \int_{\Omega_g} v_g(x, y)\psi_1(x, y)dS, \quad (33)$$

where $\Omega_g$ is the area of the vibrating surface. Then using Eqs. (19), (20), and (26), the wave component vectors $a_i$ and $b_i$ are obtained as

$$a_i = \frac{1}{2}(Z_1^{(L)} + Z_{C,1})V_1^{(L)}, \quad (34)$$

$$b_i = \frac{1}{2}(Z_1^{(L)} - Z_{C,1})V_1^{(L)}. \quad (35)$$

The wave component vectors $a_i$ and $b_i$ at each section can be calculated iteratively using Eqs. (19)–(22) and (26). Then the sound-pressure distribution can be calculated from Eq. (15).
transient sounds and investigations on the effects of the time-varying properties at the glottis [14,15]. As the presented formulation is an extension of the one-dimensional model in the frequency domain, a branch section can be represented in terms of plane waves merely by setting the number of modes at the branch section as one. In this case, ports will be added to the ideal transformer at the junction for furcating tracts and cavities. However, this treatment can be used at frequencies where transverse resonances will not appear at the branch section.

### 3. EVALUATION OF TRANSFER CHARACTERISTICS

#### 3.1. One-Dimensional Case

Figure 4 shows an electrical equivalent representation of the one-dimensional speech production model. The transfer function of the vocal tract, $H_1$, can be defined as the ratio of the volume velocity at the lips, $U_L$, to that of the glottis, $U_G$. $H_1$ can be easily calculated using a cascade matrix $F$ representing the vocal tract as an equivalent 2-port electrical circuit as follows:

$$H_1 = \frac{U_L}{U_G} = \frac{1}{C_1Z_L + D_1}.$$  \hfill (36)

$C_1$ and $D_1$ are the components of $F$, and $Z_L$ is a radiation impedance at the lips. Then a sound pressure at a far point, $P_R$, can be obtained as

$$P_R = U_L Z_T = U_G H_1 Z_T,$$  \hfill (37)

where $Z_T$ is a transfer radiation impedance and has frequency characteristics approximately +6 dB/octave, indicating high-pass properties, which corresponds to the first-order derivative of signals with respect to time.

#### 3.2. Three-Dimensional Case

The radiating area at the lips is spatially continuous, which means that a clear terminal boundary of the vocal tract is difficult to specify from the physical shape of the lips. In the vicinity of the lips, sound waves radiated from both the lips and the nostrils are superposed with different phases. As a result, extremely low sound-pressure areas may appear at some frequencies. Furthermore, at higher frequencies, higher order modes may be radiated, forming a complicated sound-pressure distribution in the vicinity of the radiating surface. Thus, the spatially averaged values of sound pressure and particle velocity in the radiating area are not appropriate for evaluating the output of the vocal tract.

3.2.1. Radiation power

Assume a sound pressure $p(r)$ and a particle velocity $v(r)$ at the position $r$. The active acoustic intensity $I(r)$ is obtained as

$$I(r) = \text{Re}\{p^*(r)v(r)/2\},$$  \hfill (38)

where $\text{Re}\{\}$ represents taking a real part, and $^*$ denotes a complex conjugate. The total acoustic power passing through an arbitrary plane $S$ in the vocal tract and a closed surface $C$ surrounding the lips and nostrils, as shown in Fig. 5, can be calculated as

$$W_S = \int_S I(r)ds \geq \int_C I(r)ds = W_C.$$  \hfill (39)

$W_C$ represents the total radiation power, and can be approximately calculated using $p(r)$ and $v(r)$ in the vicinity of the lips. In the proposed model, $W_C$ is easily calculated using the modal sound-pressure vector $P_N^{(R)}$ and velocity vector $V_N^{(R)}$ at the final section $N$, namely, on the radiating surface $C$.

$$W_C = \int_C \text{Re}\{p^* v_c/2\}ds = \frac{1}{2} \text{Re}\{\Re(P_N^{(R)}*V_N^{(R)})\}.$$  \hfill (40)

The difference between $W_S$ and $W_C$ is the power dispersion in the vocal tract when $W_S$ is evaluated at the glottis end. In the case of the lossless vocal-tract model, $W_C$ is always equal to $W_S$, which makes it easier to calculate the radiation power.

3.2.2. Evaluation by transfer impedance

Using the sound pressure $P_R$ at a far point $R$, the total radiation power $W_C$ can be written as

$$W_C = \int_C I(r)ds \approx \frac{1}{2\rho c} \int_C |P_R|^2ds.$$  \hfill (41)
Then, $|P_R|$ is obtained as

$$|P_R| = K_D \sqrt{W_C},$$  \hspace{1cm} (42)

where $K_D$ includes the directivity factor $D$ of radiation, and is defined as

$$K_D = \frac{2\rho c}{\sqrt{\int D^2 ds}}.$$  \hspace{1cm} (43)

The transfer characteristics of the model are evaluated in terms of the transfer impedance $Z_p$, the ratio of the sound pressure $P_R$ to the given source volume velocity $U_G$.

$$|Z_p| = \left| \frac{P_R}{U_G} \right| = K_D \sqrt{\frac{W_C}{|U_G|}}$$  \hspace{1cm} (44)

As the radiated power can be calculated without specifying a fixed output point, the transfer characteristics of the three-dimensional model can be evaluated using $|Z_p|$. In the following computation, $K_D$ is set to be constant although $D$ has a frequency dependence.

4. COMPUTATIONAL RESULTS

4.1. Attenuation and Phase Constants

As the wall impedance of the vocal tract is an important physical parameter, its value has been directly measured or indirectly estimated by various methodologies. The measured values showed a relatively wide range depending on the measurement methods and the position in the human body [16–21]. Here, the attenuation constant $\alpha_{mn}$ and phase constant $\beta_m$ are computed for various values of the resistive term $R$ and inertia term $L$ of the wall impedance $Z_w = R + j(\omega L - K/\omega)$. The stiffness term $K$ was ignored in this computation. The sound speed $c = 3.533 \times 10^4 \text{ cm/s}$ and air density $\rho = 1.142 \times 10^{-3} \text{ g/cm}^3$ as the values at a temperature of 36°C are used in the following computation. The vertical and horizontal (or lateral) directions are selected as the $x$- and $y$-directions, respectively.

Figure 6 shows the attenuation of modes (0, 0), (0, 1), and (0, 2) per unit length for the tube dimensions of $L_x = 1.5 \text{ cm}$ and $L_y = 5.0 \text{ cm}$. Vertical dashed lines indicate cutoff frequencies for rigid wall condition, $f_{01} = 3.539 \text{ Hz}$, $f_{02} = 2f_{01}$. Variation with respect to (a) $R$ and (b) $L$.

For higher order modes, an upward shift might be induced when the peak frequency is very close to the cutoff frequencies.

4.2. Transfer Characteristics

4.2.1. Uniform rectangular tube

The transfer characteristics of a uniform rectangular tube ($L_x = 1.5 \text{ cm}$, $L_y = 5.0 \text{ cm}$, and total length $L_z = 17 \text{ cm}$) are shown in Fig. 8 for different values of the inertia term $L$. At the entrance of the uniform tube, a square sound source (1.5 cm x 1.5 cm) is provided with a 7.25 mm offset from the left side, in order to excite higher order modes at the source boundary, as computed from Eq. (33). Modes (0, 0), (0, 1), and (0, 2) are considered in this computation. The result for the case of the rigid wall condition is also shown in the top panel for the purpose of comparison.

The bandwidth of the lower peaks, especially those below 1 kHz, are strongly influenced by the variation of $L$. Compared with the case of the rigid wall condition, a small upward shift of peak frequencies is also seen. This
plane waves. It should also be noted that even though the transfer characteristics contain only peaks arising from uniform tube, no higher order modes are excited, and the uniform vibration on the whole plane at the entrance of the junctions in the model. If the sound source is given as a general, the place of excitation is the source boundary or whether the higher order modes are excited or not. In the higher order modes depends on the condition of dimensional theory [22]. The appearance of peaks due to power was radiated than that computed from the one-tube also showed that at higher frequencies, much more measurement of the radiated power from a rectangular model with plane waves only. Experiments on the mode, much more power is radiated compared with the higher order modes in the tube become the propagative using radiated power, this result indicates that when the order modes. As the transfer characteristics are evaluated these peaks are the resonances of the propagative higher order modes decrease exponentially in the direction of the -axis, and almost vanish at the other end.

The three-dimensional shape of / measured by MRI has been converted into 36-section models. The width and height of each section are determined from the area and the maximum lateral distance of the MRI data. The entire configuration is constructed by the following procedure. First, each tube is aligned to a common horizontal plane, maintaining the symmetry with respect to the lateral direction. Then the axis position of each tube is perturbed randomly in both the vertical and lateral directions in order to imitate a small change in the vocal-tract shape. The axis position \((x_i, y_i)\) of the \(i\)th section is modified as follows, while the area of that section is kept unchanged.

result is consistent with the known effects of the wall impedance suggested by the conventional one-dimensional model.

Some sharp peaks are clearly seen above 3.5 kHz. These peaks are the resonances of the propagative higher order modes. As the transfer characteristics are evaluated using radiated power, this result indicates that when the higher order modes in the tube become the propagative mode, much more power is radiated compared with the model with plane waves only. Experiments on the measurement of the radiated power from a rectangular tube also showed that at higher frequencies, much more power was radiated than that computed from the one-dimensional theory [22]. The appearance of peaks due to the higher order modes depends on the condition of whether the higher order modes are excited or not. In general, the place of excitation is the source boundary or junctions in the model. If the sound source is given as a uniform vibration on the whole plane at the entrance of the uniform tube, no higher order modes are excited, and the transfer characteristics contain only peaks arising from plane waves. It should also be noted that even though the higher order modes are excited at the source boundary, if the modes are evanescent, only peaks of plane waves appear in the transfer characteristics as the excited higher order modes decrease exponentially in the direction of the \(z\) axis, and almost vanish at the other end.

At higher frequencies above about 3.5 kHz, the frequencies and the bandwidths of the peaks are almost the same as those under the rigid wall condition. As already seen in Fig. 6, even though the attenuation of the higher order modes is larger than that of plane waves, the wall can be regarded as rigid enough for this value of wall impedance.

4.2.2. Asymmetric multiple sections

Fig. 7 Difference in the phase constant relative to that under the rigid wall condition, \((\beta_{\text{uniform}} - \beta_{\text{rigid}})/\beta_{\text{rigid}} \times 100\%\). Variation with respect to (a) \(R\) and (b) \(L\).

Fig. 8 Transfer characteristics of uniform rectangular tube with \(L_x = 1.5\) cm, \(L_y = 5.0\) cm, and \(L_z = 17\) cm. \(Z_w = R + j\omega L, \quad R = 1,500\) g/(cm\(^2\)s). From top to bottom, rigid wall, \(L = 2.4, 1.2, 0.6, \text{and } 0.3\) g/cm\(^2\).
Here, $L_{x,i}$ and $L_{y,i}$ are the dimensions of the $i$th section and $R[-1, 1]$ is a uniform random number between $-1$ and $1$. $\varepsilon$ is a dimensionless parameter smaller than 0.5 that specifies the maximum perturbation of the axis position. Figure 9 shows the initial shape and four examples from among 100 models having slight differences in shape. $\varepsilon = 0.08$ is specified for these models.

Although the original MRI data was not for vowels, a sound source in the shape of a slit (5 mm in width) is assumed at the center of the entrance of the first section in order to show the influence of the offset connection of tubes on the transfer characteristics. As the cutoff frequencies differ in each section, five higher order modes are selected in order of frequency, and are considered in the computation together with plane waves. As a wall impedance value, the resistive and inertia terms of $R = 1.900$ g/(cm$^2$-s) and $L = 1.2$ g/cm$^2$ are used for all sections.

Figure 10 shows the transfer characteristics corresponding to the 100 models in Fig. 9. It is clearly seen that lower peaks are stable, albeit with small changes in shape. All peaks below 4 kHz arise from the resonance of plane waves. As there is no difference in the area functions for these 100 models, transfer characteristics at the lower frequencies are almost the same.

Compared with the case of the rigid wall condition (Fig. 10(a)), an increase of bandwidth and a small upward shift of the first and second peaks are also seen (Fig. 10(b)). The characteristics at higher frequencies, however, are similar, indicating less sensitivity to the values of wall impedance. A simple interpretation of this result is that the vibration of the vocal-tract wall decreases with rising frequency because of the inertia term $L$, resulting in slight attenuation of both plane waves and propagative higher order modes.

Although the lower peaks due to the plane waves are stable, there is still an influence from the higher order modes that are excited at each junction in order to satisfy the continuity of sound pressure and particle velocity at the junction, even at lower frequencies. These modes can be coupled to plane waves at the junction through the coupling matrix in Eq. (23). If the excited modes are evanescent, they never produce additional peaks at lower frequencies. However, the frequencies of peaks arising from the resonance of plane waves are influenced by the evanescent higher order modes through mode coupling, which can be observed as a slight difference in lower peak frequencies in both Figs. 10(a) and 10(b).

The transfer characteristics above 4 kHz are strongly influenced by the small change in the axis position with the appearance of sharp peaks and zeros. The frequency of the first zero is distributed in the range roughly from 4.4 kHz to 5 kHz depending on the configuration of each model. This
frequency is closely related to the offset values of the axes of the third section, having the maximum lateral dimension \( L_{x3} = 5.4 \text{ cm} \), and the adjacent sections. A possible reason for the appearance of the zeros is a transverse resonance confined in these sections as an observation of a sound-pressure distribution showed a pattern of a strong standing wave in the transverse direction in these sections. The sharp peaks at the higher frequencies arise from propagative higher order modes, as explained in Sect. 4.2.1. The excitation of the modes at each junction depends on the relative position where two tubes are joined, which is strongly influenced by the change in the axis position.

5. CONCLUSIONS

The wall impedance of the vocal tract was introduced in the computation of the propagation constant of higher order modes. The frequency characteristics of both the attenuation and phase constants, and the transfer impedance were presented. The results showed an increase in bandwidth and a small upward shift of peaks, especially at lower frequencies, while the sharp peaks at higher frequencies were less sensitive to the wall impedance.

Other internal loss factors, such as air properties (viscosity, heat conductivity, and relaxational absorption) and radiation from the skin, should be further investigated (viscosity, heat conductivity, and relaxational absorption).

ACKNOWLEDGEMENTS

The author would like to thank Dr. Hiroki Matsuzaki for valuable discussion on the acoustic analysis of the vocal tract. Part of this work has been supported by a research project of the High-Tech Research Center, Hokkai-Gakuen University, and a Grant-in-Aid for Scientific Research (No. 23300071).

REFERENCES