APPENDIX 11-A

EXAMPLES OF FEATURES FOR SOME TYPICAL VOWELS

A11-A.1 INTRODUCTION

Appendix 5 presents twelve vowels that illustrate the positions of the jaw, tongue, and lips; that is, the approximated vocal tract configuration in the form of a midsagittal outline (upper left of figure), along with the synthesized acoustic waveform (upper right), the vocal tract cross-sectional area function (lower left), and the vocal tract frequency response (lower right). These results were obtained using the articulatory speech synthesizer, which minimized the error between the vowel target formants and the articulatory model formants. The vowels and the error are listed here. Consult Appendix 5 for the other data.

- Vowel IY as in beet, error 0.01432%
- Vowel IH as in bit, 0.00632%
- Vowel EY as in bait, error 0.03832%
- Vowel EH as in bet, error 0.01204%
- Vowel AE as in bat, error 0.06795%
- Vowel UW as boot, error 0.02982%
- Vowel UH as in book, error 0.01867%
- Vowel OW as boat, error 0.19218%
- Vowel AO as in bought, error 0.05938%
- Vowel AA as in Bach, error 0.065087%
- Vowel AH as in but, error 0.21787%
- Vowel ER as in bird, error 0.57565%

APPENDIX **11–B**

ACOUSTIC TRANSFER FUNCTION CALCULATION

The formant frequencies are required by the simulated annealing algorithm to calculate the articulatory-to-acoustic inverse transform. The formant frequencies need to be decomposed from the acoustic transfer function or from the input impedance, both of which are calculated from the vocal tract cross-sectional area function. This appendix presents a derivation of the acoustic transfer function for various structures using a circuit network representation and transmission matrix theory. In Appendix 11, Section A11.2.3.1.3, it was mentioned that an elemental tube with uniform area A and length l can be analogous to the transmission line circuit model. The circuit in Figure 5.1 and the definitions given in Figure A10.27 form the basis for the following derivations. The circuit in Figure 5.1 can be rearranged as a four terminal T network, as shown in Figure A11–B.1. This network representation facilitates the acoustic transfer function calculation, which is defined as the ratio of the total output volume velocity from the vocal tract and the nasal tract to the excitation volume velocity source input. Define the propagation constant as $\gamma = \sqrt{zy}$, where $z = R + j\omega L$, $y = G + j\omega C + \frac{1}{Z_w}$, and $Z_w = R_w + j\omega L_w + \frac{1}{j\omega C_w}$. Let the characteristic impedance be

$$Z_a = Z_0 \tanh\left(\frac{\gamma \ell}{2}\right)$$

and

$$Z_b = \frac{Z_0}{\sinh(\lambda \ell)}$$

See Fant (1960) and Flanagan (1972a) for more details.

It is well known that the input-output characteristics of a four-terminal network are described by a matrix equation of the form

$$\begin{bmatrix} P_i(\omega) \\ U_i(\omega) \end{bmatrix} = [T(\omega)] \begin{bmatrix} P_o(\omega) \\ U_o(\omega) \end{bmatrix} \tag{A11–B.1}$$

where $P_i(\omega)$ and $U_i(\omega)$ are the sound pressure and volume velocity, respectively, at the input of the network; $P_o(\omega)$ and $U_o(\omega)$ are the corresponding outputs of the network; and $T(\omega)$ is the transmission matrix (also called ABCD matrix or chain matrix) of the four-terminal network. From Ohm's law and the current loop law, the transmission matrix of Figure A11–B.1 is

$$T(\omega) = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 + \frac{Z_a}{Z_b} & 2Z_a + \frac{Z_a^2}{Z_b} \\ \frac{1}{Z_b} & 1 + \frac{Z_a}{Z_b} \end{bmatrix} \tag{A11–B.2}$$

Now, consider two elemental sections with a shunt impedance Z_s in between and with radiation impedance Z_r as load. We represent the radiation impedance as $Z_r = R_r + j\omega L_r$, where $R_r = \frac{128\varrho c}{9\pi^2 A_m}$, $L_r = \frac{8\varrho}{3\pi\sqrt{\pi A_m}}$, and A_m is the lip or nostril opening area. The network representation is constructed in Figure A11–B.2. Note that the shunt impedance Z_s can be considered as the modelling circuit for the nasal sinus (see Appendix A11). Assume that the transmission matrices of section i−1 and section i

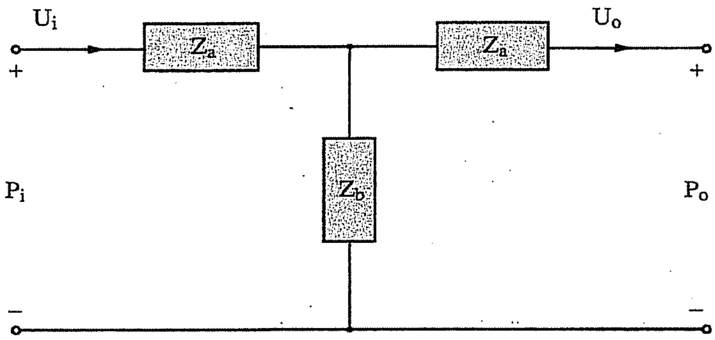


FIGURE A11–B.1 Four-terminal T-network for a uniform elemental tube.

are $T_{i-1}(\omega)$ and $T_i(\omega)$ respectively. Forming a dummy four-terminal T-network as in Figure A11–B.1 with $Z_a = 0$ and $Z_b = Z_s$ for the shunt element Z_s and applying Equation (A11–B.2), the transmission matrix is

$$T_s(\omega) = \begin{bmatrix} A_s & B_s \\ C_s & D_s \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{Z_s} & 1 \end{bmatrix} \tag{A11–B.3}$$

From the theory of cascade circuit networks, the overall transmission matrix for Figure A11–B.2 is the product of the individual transmission matrices for the sections. Thus the relation between pressure and volume velocity at the input P_{i-1}, U_{i-1} and at the radiation P_r, U_r can be written as

$$\begin{bmatrix} P_{i-1} \\ U_{i-1} \end{bmatrix} = T_{i-1}(\omega) \cdot T_s(\omega) \cdot T_i(\omega) \begin{bmatrix} P_r \\ U_r \end{bmatrix} = \begin{bmatrix} A_f & B_f \\ C_f & D_f \end{bmatrix} \begin{bmatrix} P_r \\ U_r \end{bmatrix} \tag{A11–B.4}$$

From this matrix equation and Ohm's law, $P_r = Z_r U_r$, the transfer function $H(\omega)$ from U_{i-1} to U_r is

$$H(\omega) = \frac{U_r(\omega)}{U_{i-1}(\omega)} = \frac{1}{C_f Z_r + D_f} \tag{A11–B.5}$$

and the input impedance Z_{in} is

$$Z_{in}(\omega) = \frac{P_{i-1}(\omega)}{U_{i-1}(\omega)} = \frac{A_f Z_r + B_f}{C_f Z_r + D_f} \tag{A11–B.6}$$

By use of the four-terminal network representation and the transmission matrix, we can calculate the acoustic transfer function of the vocal tract system for human speech production. Since the excitation source can be at the glottis or inside the vocal tract, there are four cases.

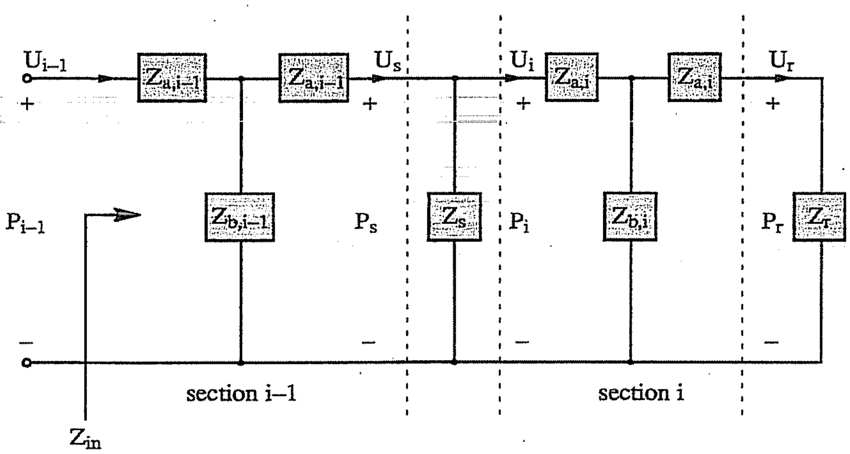


FIGURE A11–B.2 Network representation of a two-section tube with a shunt element in between and a radiation load.

Case A: Excitation Source Located at the Glottis. Let Z_{or} and Z_{nr} be the radiation impedances of the vocal tract and the nasal tract, respectively. Let Z_{sub} as the impedance of the glottis and the subglottal system looking backward from the glottis. Use $A_pB_pC_pD_p$, $A_oB_oC_oD_o$, and $A_nB_nC_nD_n$ to represent the transmission matrices of the pharyngeal tube, oral tract, and nasal tract, respectively. Figure A11–B.3(*a*) shows the equivalent network representation of the vocal system.

To calculate the acoustic transfer function $H(\omega)$ from U_g to $U_r = U_{or} + U_{nr}$, we first compute the input impedance of the oral tract and the input impedance of the nasal tract, both of which are seen downstream from the bifurcation point. From Equation (A11–B.6), we have the input impedance of oral tract as

$$Z_o = \frac{A_oZ_{or} + B_o}{C_oZ_{or} + D_o} \quad \text{(A11–B.7)}$$

and the nasal tract input impedance as

$$Z_n = \frac{A_nZ_{nr} + B_n}{C_nZ_{nr} + D_n} \quad \text{(A11–B.8)}$$

The second step is to calculate the acoustic transfer function $H_v(\omega)$ of the vocal tract from U_g to U_{or}. Forming an equivalent network for Figure A11–B.3(*a*) (see Figure A11–B.3(*b*)) and applying the analysis results of Figure A11–B.2 (refer to Equation A11–B.5), we have

$$H_v = \frac{Z_{sub}}{Z_p + Z_{sub}} \cdot \frac{1}{C_{vv}Z_{or} + D_{vv}} \quad \text{(A11–B.9)}$$

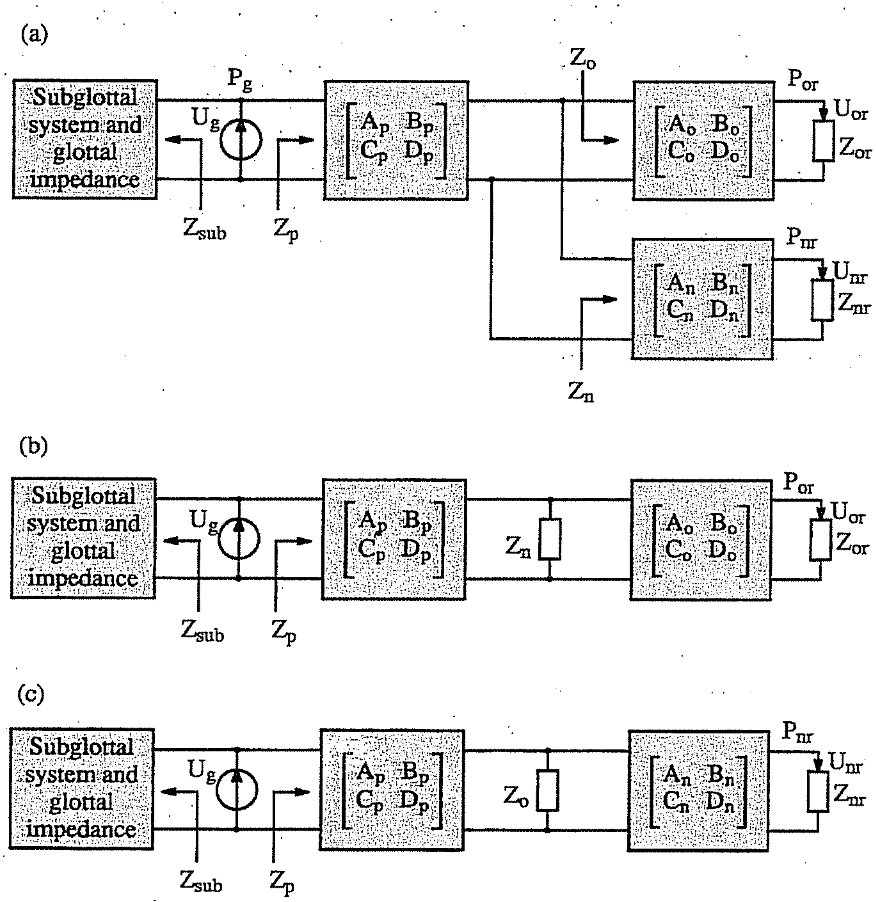


FIGURE A11–B.3 Case A with excitation source located at the glottis. (*a*) Network representation of vocal system. (*b*) Network representation for calculating $H_v(\omega)$. (*c*) Network representation for calculating $H_n(\omega)$.

and the corresponding transmission matrix as

$$\begin{bmatrix} A_{vv} & B_{vv} \\ C_{vv} & D_{vv} \end{bmatrix} = \begin{bmatrix} A_p & B_p \\ C_p & D_p \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{Z_n} & 1 \end{bmatrix} \begin{bmatrix} A_o & B_o \\ C_o & D_o \end{bmatrix} \qquad \text{(A11–B.10)}$$

Similarly, from Equation (A11–B.7), the input impedance of the vocal tract seen downstream from the glottis is

$$Z_p = \frac{A_{vv}Z_{or} + B_{vv}}{C_{vv}Z_{or} + D_{vv}} \qquad \text{(A11–B.11)}$$

Applying the same procedure to the Figure A11–B.3(c), the acoustic transfer function $H_n(\omega)$ from U_g to the output of the nasal tract U_{nr} is

$$H_n = \frac{Z_{sub}}{Z_p + Z_{sub}} \cdot \frac{1}{C_{nn}Z_{nr} + D_{nn}} \qquad \text{(A11–B.12)}$$

and the corresponding transmission matrix is

$$\begin{bmatrix} A_{nn} & B_{nn} \\ C_{nn} & D_{nn} \end{bmatrix} = \begin{bmatrix} A_p & B_p \\ C_p & D_p \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{Z_o} & 1 \end{bmatrix} \begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} \qquad \text{(A11–B.13)}$$

The vocal system acoustic transfer function is $H(\omega) = H_v(\omega) + H_n(\omega)$.

Case B: Excitation Source Located in the Pharyngeal Tube. Let $A_{pl}B_{pl}C_{pl}D_{pl}$ represent the transmission matrix of the lower part of the pharynx with the input at the excitation location and output is at the glottis, as shown in Figure A11–B.4. The results obtained in Case A can be applied to this case. Replace the Z_{sub} in Equation (A11–B.9) and (A11–B.12) with the Z'_{sub}, where the Z'_{sub} is calculated as follows.

$$Z'_{sub} = \frac{A_{pl}Z_{sub} + B_{pl}}{C_{pl}Z_{sub} + D_{pl}} \qquad \text{(A11–B.14)}$$

Then, the sum of Equations (A11–B.9) and (A11–B.12) yields the final acoustic transfer function.

Case C: Excitation Source Located at the Bifurcation of the Vocal Tract and the Nasal Tract. As shown in Figure A11–B.5, the network representation is the same as Figure A11–B.3 except the excitation source is located at the bifurcation of the vocal tract and the nasal tract. First, the input impedances Z_o and Z_n are calculated as in Equations (A11–B.7) and (A11–B.8). Then $H_v(\omega)$ and $H_n(\omega)$ are obtained as follows.

$$H_v = \frac{Z'_{sub}Z_n}{Z'_{sub}Z_o + Z_oZ_n + Z_nZ'_{sub}} \cdot \frac{1}{C_oZ_{or} + D_o} \qquad \text{(A11–B.15)}$$

$$H_n = \frac{Z'_{sub}Z_o}{Z'_{sub}Z_o + Z_oZ_n + Z_nZ'_{sub}} \cdot \frac{1}{C_nZ_{nr} + D_n} \qquad \text{(A11–B.16)}$$

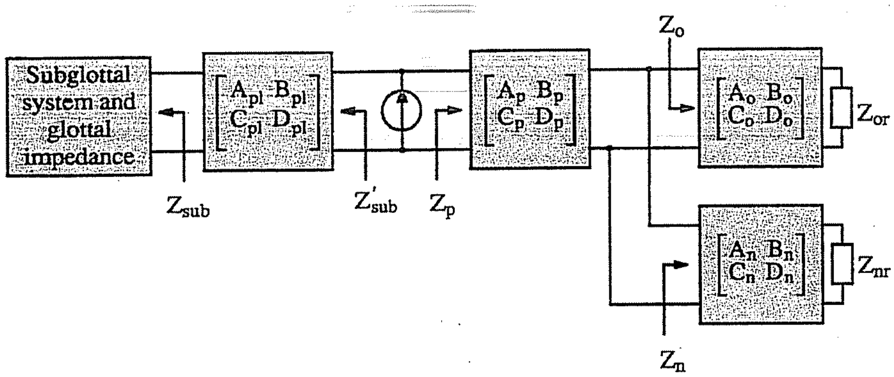


FIGURE A11–B.4 Case B with the excitation source located in the pharyngeal tube.

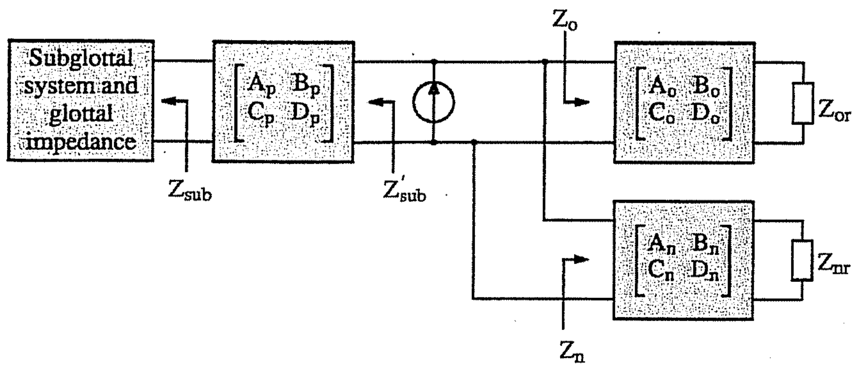


FIGURE A11-B.5 Case C with the excitation source located at the bifurcation point.

where the impedance Z'_{sub} is

$$Z'_{sub} = \frac{A_p Z_{sub} + B_p}{C_p Z_{sub} + D_p} \tag{A11–B.17}$$

The sum of $H_v(\omega)$ and $H_n(\omega)$ yields the acoustic transfer function for the vocal system.

Case D: Excitation Source Located in the Oral Tract. As shown in Figure A11–B.6(*a*), we use $A_{ob}B_{ob}C_{ob}D_{ob}$ and $A_{of}B_{of}C_{of}D_{of}$ to represent the transmission matrix of the back and front parts of oral tract, respectively. From Equations (A11–B.7), (A11–B.8), and (A11–B.17), we

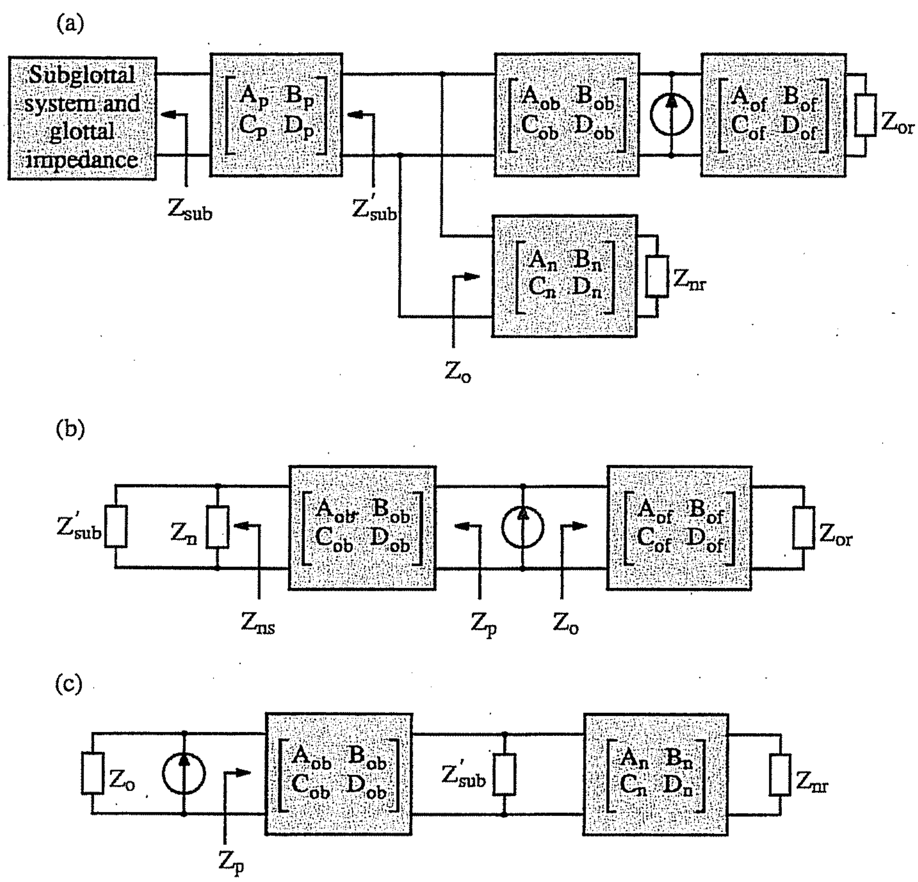


FIGURE A11-B.6 Case D with excitation source located in the oral tract. (*a*) Network representation of vocal system. (*b*) Network representation for calculating $H_v(\omega)$. (*c*) Network representation for calculating $H_n(\omega)$.

calculate the impedances Z_o, Z_n, and Z'_{sub} as follows.

$$Z_o = \frac{A_{of}Z_{or} + B_{of}}{C_{of}Z_{or} + D_{of}} \tag{A11–B.18}$$

$$Z_n = \frac{A_n Z_{nr} + B_n}{C_n Z_{nr} + D_n} \tag{A11–B.19}$$

$$Z'_{sub} = \frac{A_p Z_{sub} + B_p}{C_p Z_{sub} + D_p} \tag{A11–B.20}$$

If we define Z_{ns} (see Figure A11–B.6(*b*)) as

$$Z_{ns} = \frac{Z_n Z'_{sub}}{Z_n + Z'_{sub}} \tag{A11–B.21}$$

then the impedance Z_p can be calculated as

$$Z_p = \frac{A_{ob}Z_{ns} + B_{ob}}{C_{ob}Z_{ns} + D_{ob}} \tag{A11–B.22}$$

Now, apply Equation (A11–B.9) to obtain $H_v(\omega)$ as follows.

$$H_v = \frac{Z_p}{Z_p + Z_o} \cdot \frac{1}{C_{of}Z_{or} + D_{of}} \tag{A11–B.23}$$

For $H_n(\omega)$, we form the network as shown in Figure A11–B.6(*c*) and apply Equation (A11–B.12) to yield

$$H_n = \frac{Z_o}{Z_p + Z_o} \cdot \frac{1}{C_{on}Z_{nr} + D_{on}} \tag{A11–B.24}$$

where the corresponding transmission matrix is

$$\begin{bmatrix} A_{on} & B_{on} \\ C_{on} & D_{on} \end{bmatrix} = \begin{bmatrix} A_{ob} & B_{ob} \\ C_{ob} & D_{ob} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{Z'_{sub}} & 1 \end{bmatrix} \begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} \tag{A11–B.25}$$

Then, the final acoustic transfer function is $H_v(\omega) + H_n(\omega)$.

We have considered the acoustic transfer function of the soft-wall, lossy vocal system. For some cases, calculation of the acoustic transfer function of a lossless or lossy, hard-wall vocal system is adequate and also easier. For a rigid and lossy vocal system, the wall impedance Z_w of each section is removed, causing the impedance Z_w to be infinity. In this case, the admittance y becomes $y = G + j\omega C$. For a rigid and lossless vocal system, this means that no dissipation elements and no wall impedance are present for each section. Then the propagation constant γ becomes an imaginary quantity.

In the above derivations, we assume that impedance Z_{sub} is composed of the Foster-chain circuit model for the subglottal system with glottal impedance $Z_g = R_g + j\omega L_g$, where R_g and L_g are defined in Appendix A11. Since the glottal area is a time-varying function, we assume that the glottal area and glottal volume velocity are constants when we evaluate the acoustic transfer function. In Case A, if the subglottal system and the glottal impedance are not included, then we let the impedance $Z_{sub} \to \infty$, which corresponds to an open circuit, that is, the block representing the subglottal system and the glottal impedance in Figure A11–B.3 are disconnected. For other cases, two glottal conditions are considered. If the glottis is open, the effects of the subglottal system and the glottal impedance are evaluated. Otherwise, if the glottis is closed, we let $Z_{sub} = 0$, that is, Z_{sub} is a short circuit. The velopharyngeal port opening area, which couples the nasal tract to the vocal system, also affects the acoustic transfer function. Our software system allows the user to change the velopharyngeal port opening area, the number of coupled sinus cavities, and the location of the excitation source when evaluating the acoustic transfer function.

APPENDIX **11–C**

DERIVATION OF DISCRETE TIME ACOUSTIC EQUATIONS

As covered in Section A11.2.3.1, the vocal tract tube can be described by two coupled partial differential acoustic equations. These two acoustic equations are functions of both time and space. Approximating the vocal tract as a sequence of elemental sections corresponds to digitizing the vocal tract in space, that is, spatial sampling. For each elemental section, the transmission-line analog approach is applied to form the equivalent circuit model, as seen in Figure 5.1. Connecting the equivalent circuit of each section together in combination with the equivalent circuit models of the other parts of the vocal system (subglottal system, glottis, and nasal sinus cavities), a lumped circuit network representation of the vocal system can be formed, as shown in Figure A11.18; the time-domain approach, the Kirchoff's and Ohm's laws are applied to the circuit network to obtain sets of differential equations. These differential equations, which correspond to the equivalent acoustic equations, which govern the generation and the propagation of acoustic waves inside the vocal system, are transformed into discrete-time representations. This appendix provides a detailed derivation of the discrete-time acoustic equations, that is, the difference matrix equations. The discretization scheme is similar to the work of Maeda (1982a). Our model, however, provides more features, such as the subglottal system, nasal sinus cavities, and turbulence noise source.

Consider the transmission-line circuit model, shown in Figure A11–C.1, of the ith section of the vocal tract. Define the volume velocity (current) at the input of the ith section, that is, at x_{i-1} in space, as $U_i(t) = u_i(x_{i-1}, t)$, where $x_{i-1} = \sum_{k=1}^{i-1} l_k$ is the vocal tract length from the glottis to the ith section. Similarly, define the central pressure (voltage) of the ith section as $P_i(t) = p_i(x_i - \frac{l_i}{2}, t)$, where l_i is the section length of section i. From Kirchoff's and Ohm's laws, we have the following differential equations.

$$
\begin{aligned}
P_{i-1}(t) - P_i(t) &= \frac{d}{dt}\{[L_{i-1}(t) + L_i(t)]U_i(t)\} + [R_{i-1}(t) + R_i(t)]U_i(t) \\
U_i(t) - U_{i+1}(t) &= u_{i1}(t) + u_{i2}(t) + u_{i3}(t) \\
&= G_i(t)P_i(t) + \frac{d}{dt}\{C_i(t)P_i(t)\} + u_{i3}(t) \\
P_i(t) &= R_{w,i}(t)u_{i3}(t) + \frac{d}{dt}\{L_{w,i}(t)u_{i3}(t)\} + \int_0^t \left\{\frac{u_{i3}(\tau)}{C_{w,i}(\tau)}\right\} d\tau
\end{aligned}
\qquad \text{(A11–C.1)}
$$

There are three terms in Equation (A11–C.1); simple, differential, and integral terms. Define the general forms of these three terms as follows.

$$
\begin{aligned}
y_1(t) &\equiv c_1(t)x(t) \\
y_2(t) &\equiv \frac{d}{dt}\{c_2(t)x(t)\} \\
y_3(t) &\equiv \int_0^t \{c_3(\tau)x(\tau)\}\, d\tau
\end{aligned}
\qquad \text{(A11–C.2)}
$$

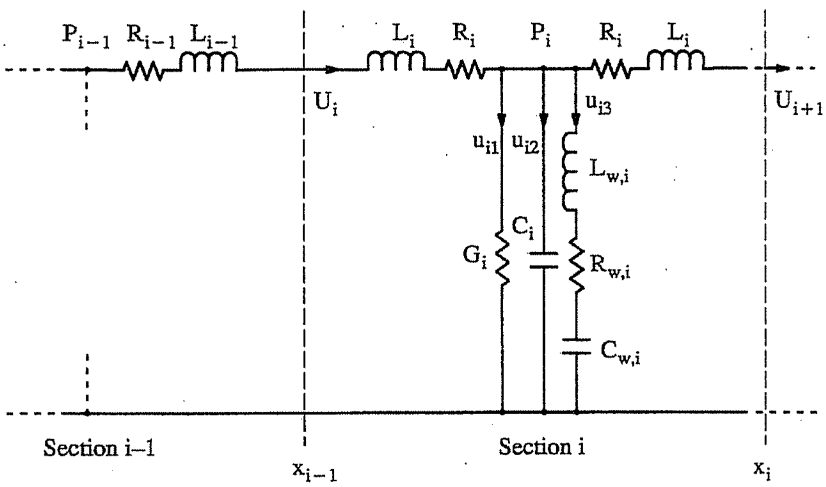


FIGURE A11–C.1 A lumped transmission-line circuit model of the ith section.

where $c_i(t)$ ($i = 1$, 2, or 3) is a coefficient that represents the time-varying circuit component, and $x(t)$ can be $P_i(t)$, $U_i(t)$, or $u_{i3}(t)$. Let $y_i(n) = y_i(t = nT)$, $c_i(n) = c_i(t = nT)$, and $x(n) = x(t = nT)$ represent the sampled values of $y_i(t)$, $c_i(t)$, and $x(t)$, respectively, at $t = nT$, where $n = 0, 1, 2, \ldots$, and T denotes the sampling time interval. For the first term, it is obvious that

$$y_1(n) = c_1(n)x(n) \qquad \text{(A11–C.3)}$$

For the differential term, we first integrate both sides

$$\int_{(n-1)T}^{nT} y_2(t)\,dt = \int_{(n-1)T}^{nT} \frac{d}{dt}\{c_2(t)x(t)\}\,dt = \int_{(n-1)T}^{nT} d\{c_2(t)x(t)\} \qquad \text{(A11–C.4)}$$

The trapezoidal rule is applied to the left hand side of Equation (A11–C.4) to yield

$$\int_{(n-1)T}^{nT} y_2(t)\,dt \equiv \frac{T}{2}[y_2(n) + y_2(n-1)] \qquad \text{(A11–C.5)}$$

Then the discrete-time form of Equation (A11–C.4) is

$$\frac{T}{2}[y_2(n) + y_2(n-1)] = c_2(n)x(n) - c_2(n-1)x(n-1) \qquad \text{(A11–C.6)}$$

Equation (A11–C.6) can be rewritten in recursive form as

$$y_2(n) = \frac{2}{T}c_2(n)x(n) - Q(n-1) \qquad \text{(A11–C.7)}$$

where

$$Q(n-1) = \frac{4}{T}c_2(n-1)x(n-1) - Q(n-2)$$

Similarly, the integral term is approximated by

$$y_3(n) = \frac{T}{2}c_3(n)x(n) + V(n-1) \qquad \text{(A11–C.8)}$$

where

$$V(n-1) = Tc_3(n-1)x(n-1) + V(n-2)$$

Now applying the rules of Equations (A11–C.3), (A11–C.7), and (A11–C.8) to the set of differential Equations (A11–C.1), the transformed equations are

$$P_{i-1}(n) - P_i(n) = \left\{R_{i-1}(n) + R_i(n) + \frac{2}{T}[L_{i-1}(n) + L_i(n)]\right\} U_i(n) - Q_{L_{i-1}L_i}(n-1)$$

$$U_i(n) - U_{i+1}(n) = \left\{G_i(n) + \frac{2}{T}C_i(n)\right\} P_i(n) - Q_{C_i}(n-1) + u_{i3}(n) \qquad \text{(A11–C.9)}$$

$$P_i(n) = \left\{R_{w,i}(n) + \frac{2}{T}L_{w,i}(n) + \frac{T}{2}\frac{1}{C_{w,i}(n)}\right\} u_{i3}(n) + V_{C_{w,i}}(n-1) - Q_{L_{w,i}}(n-1)$$

where

$$Q_{L_{i-1}L_i}(n-1) = \frac{4}{T}[L_{i-1}(n-1) + L_i(n-1)]U_i(n-1) - Q_{L_{i-1}L_i}(n-2)$$
$$Q_{C_i}(n-1) = \frac{4}{T}C_i(n-1)P_i(n-1) - Q_{C_i}(n-2)$$
$$V_{C_{w,i}}(n-1) = \frac{T}{C_{w,i}(n-1)}u_{i3}(n-1) + V_{C_{w,i}}(n-2)$$
$$Q_{L_{w,i}}(n-1) = \frac{4}{T}L_{w,i}(n-1)u_{i3}(n-1) - Q_{L_{w,i}}(n-2)$$

Define

$$\begin{aligned}
a_i(n) &= \frac{2}{T}L_i(n) + R_i(n) \\
Y_{w,i}(n) &= \frac{1}{\frac{2}{T}L_{w,i}(n) + R_{w,i}(n) + \frac{T}{2}\frac{1}{C_{w,i}(n)}} \\
b_i(n) &= \frac{1}{\frac{2}{T}C_i(n) + G_i(n) + Y_{w,i}(n)} \\
V_i(n-1) &= Q_{C_i}(n-1) - Y_{w,i}(n)[Q_{L_{w,i}}(n-1) - V_{C_{w,i}}(n-1)] \\
u_{i3}(n) &= Y_{w,i}(n)[P_i(n) + Q_{L_{w,i}}(n-1) - V_{C_{w,i}}(n-1)] \\
H_i(n) &= a_{i-1}(n) + a_i(n) + b_{i-1}(n) + b_i(n) \\
F_i(n) &\equiv b_{i-1}(n)V_{i-1}(n-1) + Q_{L_{i-1}L_i}(n-1) - b_i(n)V_i(n-1)
\end{aligned} \tag{A11–C.10}$$

After some manipulations, we obtain the difference equations for the ith section, as follows.

$$\begin{aligned}
F_i(n) &= -b_{i-1}(n)U_{i-1}(n) + H_i(n)U_i(n) - b_i(n)U_{i+1}(n) \\
P_i(n) &= b_i(n)[U_i(n) - U_{i+1}(n) + V_i(n-1)]
\end{aligned} \tag{A11–C.11}$$

Next, consider a shunt element, which models the nasal sinus, inserted in between two sections, i and i + 1. Figure A11–C.2 is the corresponding circuit model. The set of differential equations is

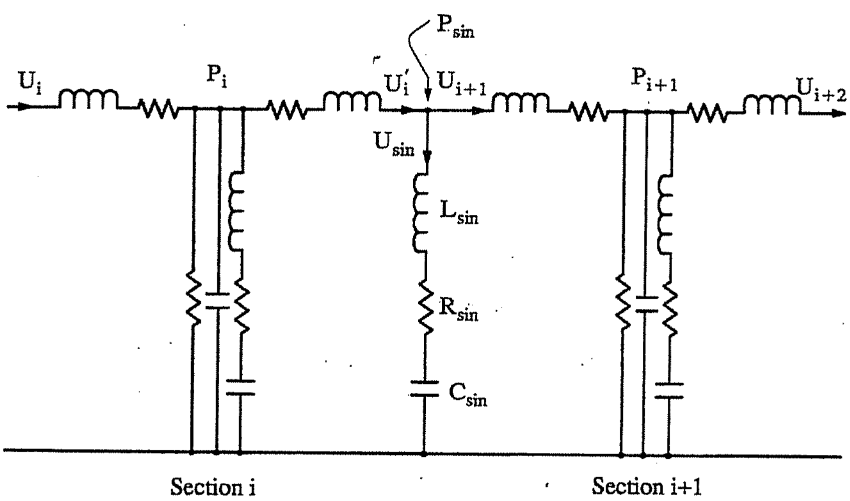


FIGURE A11–C.2 A shunt element of the nasal sinus inserted between two sections.

given by

$$\begin{aligned} P_i(t) - P_{sin}(t) &= \frac{d}{dt}\left[L_i(t)U_i'(t)\right] + R_i(t)U_i'(t) \\ P_{sin}(t) - P_{i+1}(t) &= \frac{d}{dt}[L_{i+1}(t)U_{i+1}] + R_{i+1}(t)U_{i+1}(t) \\ P_{sin}(t) &= R_{sin}U_{sin}(t) + L_{sin}\frac{dU_{sin}(t)}{dt} + \frac{1}{C_{sin}}\int_0^t U_{sin}(\tau)\,d\tau \end{aligned} \tag{A11–C.12}$$

Applying the rules of Equations (A11–C.3), (A11–C.7), and (A11–C.8) to Equation set (A11–C.12), the corresponding difference equations are

$$\begin{aligned} P_i(n) - P_{sin}(n) &= a_i(n)U_i'(n) - Q_{L_i}'(n-1) \\ P_{sin}(n) - P_{i+1}(n) &= a_{i+1}(n)U_{i+1}(n) - Q_{L_{i+1}}(n-1) \\ P_{sin}(n) &= b_{sin}\left[U_i'(n) - U_{i+1}(n)\right] - Q_{L_{sin}}(n-1) + V_{C_{sin}}(n-1) \end{aligned} \tag{A11–C.13}$$

where

$$\begin{aligned} a_i(n) &= \frac{2}{T}L_i(n) + R_i(n) \\ U_{sin}(n) &= U_i'(n) - U_{i+1}(n) \\ Q_{L_i}'(n-1) &= \frac{4}{T}L_i(n-1)U_i'(n-1) - Q_{L_i}'(n-2) \\ Q_{L_{i+1}}(n-1) &= \frac{4}{T}L_{i+1}(n-1)U_{i+1}(n-1) - Q_{L_{i+1}}(n-2) \\ Q_{L_{sin}}(n-1) &= \frac{4}{T}L_{sin}U_{sin}(n-1) - Q_{L_{sin}}(n-2) \\ V_{C_{sin}}(n-1) &= \frac{4}{C_{sin}}U_{sin}(n-1) + V_{C_{sin}}(n-2) \\ b_{sin} &= \left[\frac{2}{T}L_{sin} + R_{sin} + \frac{T}{2C_{sin}}\right] \end{aligned}$$

After rearranging items, the difference equations governing the shunt element can be rewritten as

$$\begin{aligned} F_{sin}(n) &= -b_i(n)U_i(n) + H_{sf}(n)U_i'(n) - b_{sin}U_{i+1}(n) \\ F_{i+1}(n) &= -b_{sin}U_i'(n) + H_{sb}(n)U_{i+1}(n) - b_{i+1}(n)U_{i+2}(n) \\ P_{sin}(n) &= b_{sin}\left[U_i'(n) - U_{i+1}(n)\right] - Q_{L_{sin}}(n-1) + V_{C_{sin}}(n-1) \\ P_i(n) &= b_i(n)\left[U_i(n) - U_i'(n) + V_i(n-1)\right] \end{aligned} \tag{A11–C.14}$$

where

$$\begin{aligned} F_{sin}(n) &\equiv Q_{L_{sin}}(n-1) + Q_{L_i}'(n-1) + b_i(n)V_i(n-1) - V_{C_{sin}}(n-1) \\ F_{i+1}(n) &\equiv Q_{L_{i+1}}(n-1) - Q_{L_{sin}}(n-1) - b_{i+1}(n)V_{i+1}(n-1) + V_{C_{sin}}(n-1) \\ H_{sf}(n) &= a_i(n) + b_i(n) + b_{sin} \\ H_{sb}(n) &= a_{i+1}(n) + b_{i+1}(n) + b_{sin} \end{aligned}$$

Now, we consider the circuit representation at the bifurcation point of the vocal tract and the nasal tract. From Figure A11–C.3, the differential equations are given by

$$\begin{aligned} P_i(t) - P_{NC}(t) &= \frac{d}{dt}[L_i(t)U_{NC}(t)] + R_i(t)U_{NC}(t) \\ P_{NC}(t) - P_{i+1}(t) &= \frac{d}{dt}[L_{i+1}(t)U_{i+1}] + R_{i+1}(t)U_{i+1}(t) \\ P_{NC}(t) - P_{N1}(t) &= \frac{d}{dt}[L_{N1}(t)U_{N1}] + R_{N1}(t)U_{N1}(t) \\ U_i(t) - U_{NC}(t) &= u_{i1}(t) + u_{i2}(t) + u_{i3}(t) \\ &= G_i(t)P_i(t) + \frac{d}{dt}\{G_i(t)P_i(t)\} + u_{i3}(t) \\ U_{NC}(t) &= U_{i+1}(t) + U_{N1}(t) \end{aligned} \tag{A11–C.15}$$

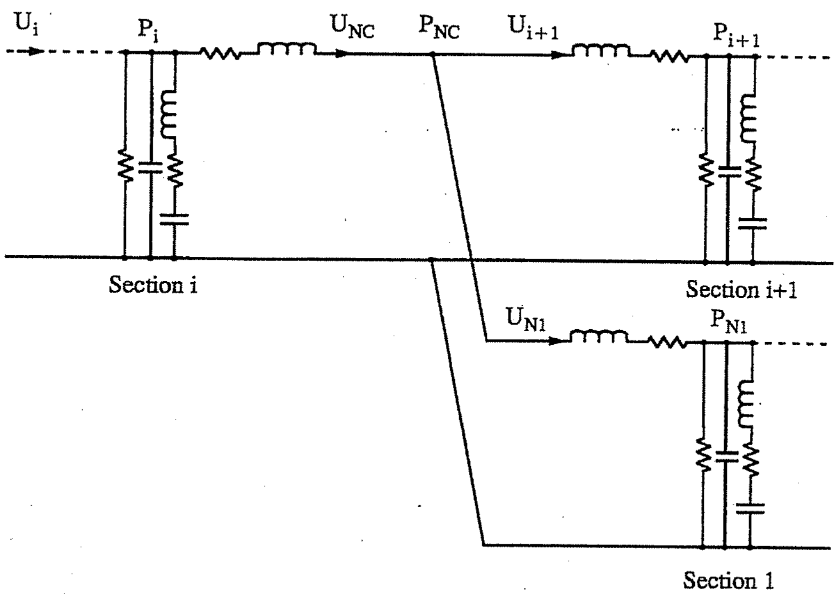


FIGURE A11–C.3 The circuit model of the bifurcation point.

After applying the discretization rules, previous derivation results, and some simple manipulations, one obtains the following difference equations

$$\begin{aligned}
F_i(n) &= -b_{i-1}(n)U_{i-1}(n) + H_i(n)U_i(n) - b_i(n)U_{NC}(n) \\
F_{NC}(n) &= -b_i(n)U_i(n) + H_{NC}(n)U_{NC}(n) + P_{NC}(n) \\
F_{i+1}(n) &= -P_{NC}(n) + H_{i+1}(n)U_{i+1}(n) - b_{i+1}(n)U_{i+2}(n) \\
P_i(n) &= b_i(n)\,[U_i(n) - U_{NC}(n) + V_i(n-1)] \\
F_{N1}(n) &= -P_{NC}(n) + H_{N1}(n)U_{N1}(n) - b_{N1}(n)U_{N2}(n) \\
U_{NC}(n) &= U_{i+1}(n) + U_{N1}(n)
\end{aligned} \qquad \text{(A11–C.16)}$$

where

$$\begin{aligned}
F_i(n) &\equiv Q_{L_{i-1}L_i}(n-1) + b_{i-1}(n)V_{i-1}(n-1) - b_i(n)V_i(n-1) \\
F_{NC}(n) &\equiv Q'_{L_i}(n-1) + b_i(n)V_i(n-1) \\
F_{i+1}(n) &\equiv Q_{L_{i+1}}(n-1) - b_{i+1}(n)V_{i+1}(n-1) \\
F_{N1}(n) &\equiv Q_{L_{N1}}(n-1) - b_{N1}(n)V_{N1}(n-1) \\
H_{NC}(n) &= a_i(n) + b_i(n) \\
H_{i+1}(n) &= a_{i+1}(n) + b_{i+1}(n) \\
H_{N1}(n) &= a_{N1}(n) + b_{N1}(n) \\
Q'_{L_i}(n-1) &= \frac{4}{T}L_i(n-1)U_{NC}(n-1) - Q'_{L_i}(n-2)
\end{aligned}$$

Next, we consider the radiation part. Figure A11–C.4 shows the circuit model of the mouth section and radiation. The differential equations for this circuit are

$$\begin{aligned}
P_i(t) - P_r(t) &= \frac{d}{dt}[L_i(t)U_r(t)] + R_i(t)U_r(t) \\
U_r(t) &= \frac{P_r(t)}{R_r(t)} + \int_0^t \frac{P_r(\tau)}{L_r(\tau)}\,d\tau
\end{aligned} \qquad \text{(A11–C.17)}$$

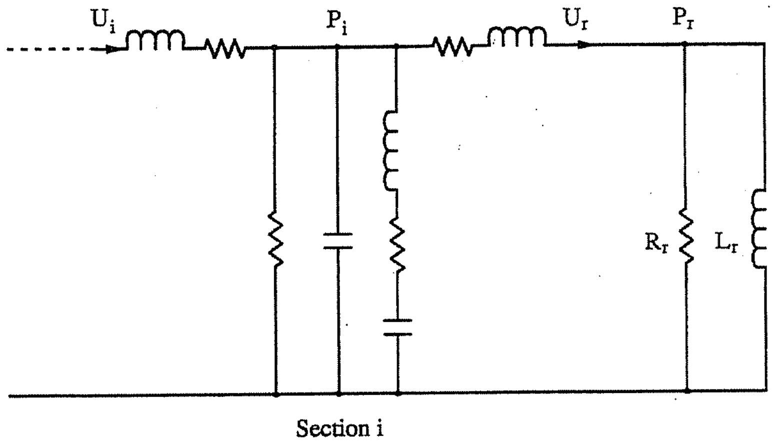


FIGURE A11–C.4 The circuit model of radiation.

The corresponding difference equations are

$$
\begin{aligned}
F_i(n) &= -b_{i-1}(n)U_{i-1}(n) + H_i(n)U_i(n) - b_i(n)U_r(n) \\
F_r(n) &= -b_i(n)U_i(n) + H_r(n)U_r(n) \\
P_i(n) &= b_i(n)\left[U_i(n) - U_r(n) + V_i(n-1)\right] \\
P_r(n) &= b_r(n)\left[U_r(n) - V_{L_r}(n-1)\right]
\end{aligned} \qquad \text{(A11–C.18)}
$$

where

$$
\begin{aligned}
F_i(n) &= Q_{L_{i-1}L_i}(n-1) + b_{i-1}(n)V_{i-1}(n-1) - b_i(n)V_i(n-1) \\
F_r(n) &= Q'_{L_i}(n-1) + b_i(n)V_i(n-1) + b_r(n)V_{L_r}(n-1) \\
H_r(n) &= a_i(n) + b_i(n) + b_r(n) \\
b_r(n) &= \frac{1}{\frac{1}{R_r(n)} + \frac{T}{2}\frac{1}{L_r(n)}} \\
Q'_{L_i}(n-1) &= \frac{4}{T}L_i(n-1)U_r(n-1) - Q_{L_i}(n-2) \\
V_{L_r}(n-1) &= \frac{T}{L_r(n-1)}P_r(n-1) + V_{L_r}(n-2)
\end{aligned}
$$

Finally, we consider the excitation input. We have two types: one is without the glottal impedance and the subglottal system; one is with the glottal impedance and the subglottal system. We first derive the difference equations for no glottal impedance and no subglottal system. From Figure A11–C.5, it is easy to obtain the difference equations from previous derivation results.

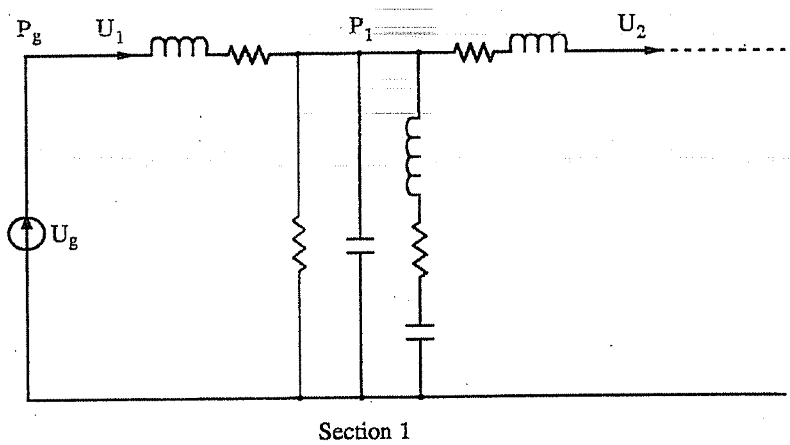


FIGURE A11–C.5 The circuit model of the first section of vocal tract with excitation.

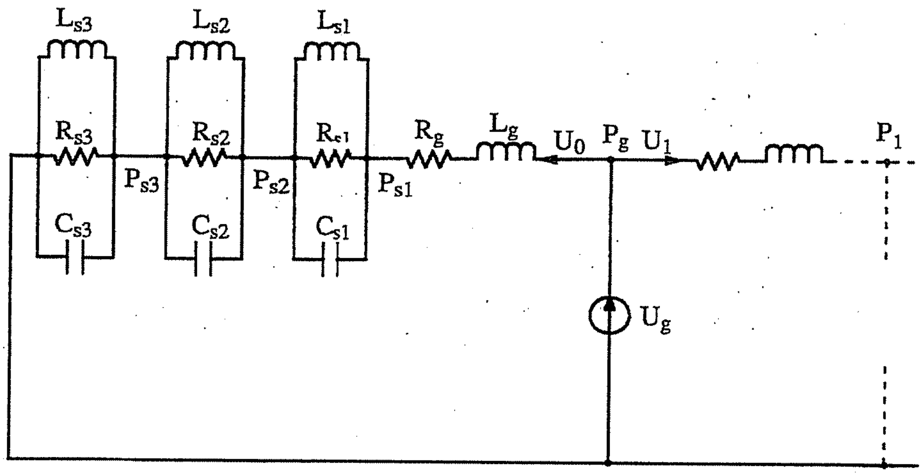


FIGURE A11–C.6 The circuit model of subglottal system and glottal impedance with excitation.

They are

$$\begin{aligned} F_1(n) &= -P_g(n) - b_1(n)U_2(n) \\ U_1(n) &= U_g(n) \\ P_1(n) &= b_1(n)\,[U_1(n) - U_2(n) + V_1(n-1)] \end{aligned} \qquad \text{(A11–C.19)}$$

where

$$F_1(n) \equiv Q_{L_1}(n-1) + b_1(n)V_1(n-1) - [a_1(n) + b_1(n)]\,U_1(n)$$

For the case of excitation source with the glottal impedance and the subglottal system, we form the circuit model in Figure A11–C.6. The corresponding differential equations are given by

$$\begin{aligned} U_0(t) &= \frac{[P_{s1}(t) - P_{s2}(t)]}{R_{s1}} + \frac{1}{L_{s1}}\int_0^t [P_{s1}(\tau) - P_{s2}(\tau)]\,d\tau + C_{s1}\frac{d}{dt}[P_{s1}(t) - P_{s2}(t)] \\ &= \frac{[P_{s2}(t) - P_{s3}(t)]}{R_{s2}} + \frac{1}{L_{s2}}\int_0^t [P_{s2}(\tau) - P_{s3}(\tau)]\,d\tau + C_{s2}\frac{d}{dt}[P_{s2}(t) - P_{s3}(t)] \\ &= \frac{P_{s3}(t)}{R_{s3}} + \frac{1}{L_{s3}}\int_0^t P_{s3}(\tau)\,d\tau + C_{s3}\frac{dP_{s3}(t)}{dt} \end{aligned} \qquad \text{(A11–C.20)}$$

$$P_g(t) - P_{s1}(t) = R_g(t)U_0(t) + \frac{d}{dt}[L_g(t)U_0(t)]$$

$$U_g(t) = U_0(t) + U_1(t)$$

Applying the discretization rules to above equations and solving $P_{si}(n)$, $i = 1, 2, 3$, in order, then the difference equations can be written as

$$\begin{aligned} F_s(n) &= P_g(n) + X_{sub}(n)U_1(n) \\ U_g(n) &= U_1(n) + U_0(n) \\ F_1(n) &= -P_g(n) + H_1(n)U_1(n) - b_1(n)U_2(n) \end{aligned} \qquad \text{(A11–C.21)}$$

where

$$F_s(n) \equiv V_{s1}(n-1) + V_{s2}(n-1) + V_{s3}(n-1) + X_{sub}(n)U_g(n) - Q_{L_s}(n-1)$$

$$F_1(n) \equiv Q_{L_1}(n-1) - b_1(n)V_1(n-1)$$

$$V_{si}(n-1) = b_{si}\left[C_{si}Q_{C_{si}}(n-1) - \frac{V_{L_{si}}(n-1)}{L_{si}}\right], \quad i = 1, 2, 3$$

$$
\begin{aligned}
b_{si}(n) &= \frac{1}{\frac{1}{R_{si}(n)} + \frac{T}{2}\frac{1}{L_{si}(n)} + \frac{2}{T}C_{si}}, \quad i = 1, 2, 3 \\
Q_{C_{si}}(n-1) &= \frac{4}{T}[P_{si}(n-1) - P_{si+1}(n-1)] - Q_{C_{si}}(n-2), \quad i = 1, 2 \\
Q_{C_{s3}}(n-1) &= \frac{4}{T}P_{s3}(n-1) - Q_{C_{s3}}(n-2) \\
V_{L_{si}}(n-1) &= T[P_{si}(n-1) - P_{si+1}(n-1)] + V_{L_{si}}(n-2), \quad i = 1, 2 \\
V_{L_{s3}}(n-1) &= TP_{s3}(n-1) + V_{L_{s3}}(n-2) \\
Q_{C_{si}}(n-1) &= \frac{4}{T}[P_{si}(n-1) - P_{si+1}(n-1)] - Q_{C_{si}}(n-2), \quad i = 1, 2 \\
Q_{C_{si}}(n-1) &= \frac{4}{T}P_{s3}(n-1) - Q_{C_{s3}}(n-2) \\
P_{si}(n) &= \sum_{k=i}^{1} b_{sk}U_0(n) + \sum_{k=i}^{1} V_{sk}(n-1), \quad i = 1, 2, 3 \\
P_g(n) &= V_{s1}(n-1) + V_{s2}(n-1) + V_{s3}(n-1) + X_{sub}(n)U_0(n) - Q_{L_g}(n-1) \\
Q_{L_g}(n-1) &= \frac{4}{T}[L_g(n-1)U_0(n-1)] - Q_{L_g}(n-2) \\
X_{sub}(n) &= a_g(n) + b_{sl} + b_{s2} + b_{s3} \\
a_g(n) &= R_g(n) + \frac{2}{T}L_g(n)
\end{aligned}
$$

We have derived the difference equations for different parts of the vocal system. Note that in order to set all recursive equations, Q(n−1) and V(n−1), in motion, the initial rest conditions of the vocal system need to be assumed; that is, Q(0) = 0 and V(0) = 0 for all sections. To write out the matrix equations for the entire vocal system, we assume that

1. The vocal tract has a number of sections denoted as NING.
2. The nasal tract has a number of sections denoted as NTN.
3. The bifurcation point is located at the downstream of a vocal tract section denoted as NTS.
4. The excitation source is located at the input to first section of vocal tract.
5. One sinus is inserted at the downstream of a nasal-tract section denoted as NS.

For the vocal system that has no glottal impedance and no subglottal system but has one nasal sinus coupling, the three sets of simultaneous difference equations or the matrix equations for the pharyngeal, oral, and nasal tracts, respectively, are

$$
\begin{aligned}
& U_1(n) = U_g(n) \\
k = 0:\ & F_1(n) = Q_{L_1}(n-1) - b_1(n)V_1(n-1) - H_1(n)U_1(n) \\
& = -P_g(n) - b_1(n)U_2(n) \\
k = 1:\ & F_2(n) = \overline{b_1(n)}V_1(n-1) + Q_{L_1L_2}(n-1) - b_2(n)V_2(n-1) + b_1(n)U_1(n) \\
& = H_2(n)U_2(n) - b_2(n)U_3(n) \\
2 \le k < NTS:\ & F_{k+1}(n) = \overline{b_k(n)}V_k(n-1) + Q_{L_kL_{k+1}}(n-1) - b_{k+1}(n)\overline{V_{k+1}}(n-1) \\
& = -b_k(n)U_k(n) + H_{k+1}(n)U_{k+1}(n) - b_{k+1}(n)U_{k+2}(n) \\
& \text{if } k+2 = NTS+1 \Rightarrow U_{k+2}(n) = U_{NC}(n) \\
k = NTS:\ & F_{NC}(n) = Q'_{L_{NTS}}(n-1) + b_{NTS}(n)V_{NTS}(n-1) \\
& = -b_{NTS}(n)U_{NTS}(n) + H_{NC}(n)U_{NC}(n) + P_{NC}(n) \\
& P_i(n) = b_i(n)[U_i(n) - U_{i+1}(n) + V_i(n)], \quad \text{where } i = 1, \ldots, NTS
\end{aligned}
$$

if

$$i = NTS,$$

then

$$U_{NTS+1}(n) = U_{NC}(n)$$

$$\begin{bmatrix} F_1 \\ F_2 \\ \cdot \\ \cdot \\ \cdot \\ F_{NTS} \\ F_{NC} \end{bmatrix} = \begin{bmatrix} -1 & -b_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & H_2 & -b_2 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & H_{NTS} & -b_{NTS} & 0 \\ 0 & 0 & 0 & \cdots & -b_{NTS} & H_{NC} & 1 \end{bmatrix} \begin{bmatrix} P_g \\ U_2 \\ \cdot \\ \cdot \\ \cdot \\ U_{NC} \\ P_{NC} \end{bmatrix} \qquad \text{(A11–C.22)}$$

$$\begin{aligned} F_{NTS+1}(n) &\equiv Q_{L_{NTS+1}}(n-1) - b_{NTS+1}(n)V_{NTS+1}(n-1) \\ &= -P_{NC}(n) + H_{NTS+1}(n)U_{NTS+1}(n) - b_{NTS+1}(n)U_{NTS+2}(n) \\ F_i(n) &\equiv Q_{L_{i-1}L_i}(n-1) + b_{i-1}(n)V_{i-1}(n-1) - b_i(n)V_i(n-1) \\ &= -b_{i-1}(n)U_{i-1}(n) + H_i(n)U_i(n) - b_i(n)U_{i+1}(n) \end{aligned}$$

where

$$i = NTS+2, \ldots, NING, \quad \text{if } i = NING$$

then

$$\begin{aligned} U_{NING+1}(n) &= U_r(n) \\ F_r(n) &\equiv b_{NING}(n)V_{NING}(n-1) + b_r(n)V_{L_r}(n-1) + Q'_{NING}(n-1) \\ &= -b_{NING}(n)U_{NING}(n) + H_r(n)U_r(n) \\ P_i(n) &= b_i(n)[U_i(n) - U_{i+1}(n) + V_i(n-1)], \quad \text{where } i = NTS+1, \ldots, NING-1 \\ P_{NING}(n) &= b_{NING}(n)[U_{NING}(n) - U_r(n) + V_{NING}(n-1)] \\ P_r(n) &= b_r(n)[U_r(n) - V_{L_r}(n-1)] \end{aligned}$$

$$\begin{bmatrix} F_{NTS+1} \\ F_{NTS+2} \\ F_{NTS+3} \\ \cdot \\ \cdot \\ \cdot \\ F_{NING} \\ F_r \end{bmatrix} = \begin{bmatrix} -1 & H_{NTS+1} & -b_{NTS+1} & 0 & \cdots & 0 & 0 & 0 \\ 0 & -b_{NTS+1} & H_{NTS+2} & -b_{NTS+2} & \cdots & 0 & 0 & 0 \\ 0 & 0 & -b_{NTS+2} & H_{NTS+3} & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & -b_{NING-1} & H_{NING} & -b_{NING} \\ 0 & 0 & 0 & 0 & \cdots & 0 & -b_{NING} & H_r \end{bmatrix} \begin{bmatrix} P_{NC} \\ U_{NTS+1} \\ U_{NTS+2} \\ \cdot \\ \cdot \\ \cdot \\ U_{NING} \\ U_r \end{bmatrix}$$

(A11–C.23)

$$\begin{aligned} F_{N1}(n) &= Q_{L_{N1}}(n-1) - b_{N1}(n)V_{N1}(n-1) \\ &= -P_{NC}(n) + H_{N1}(n)U_{N1}(n) - b_{N1}(n)U_{N2}(n) \\ F_{Ni}(n) &= Q_{L_{Ni-1}L_{Ni}}(n-1) + b_{Ni-1}(n)V_{Ni-1}(n-1) - b_{Ni}(n)V_{Ni}(n-1) \\ &= -b_{Ni-1}(n)U_{Ni-1}(n) + H_{Ni}(n)U_{Ni}(n) - b_{Ni}(n)U_{Ni+1}(n) \\ i &= 2, \ldots, N_s, N_{s+2}, \ldots, NTN; \end{aligned}$$

if

$$i = N_s$$

then

$$\begin{aligned} U_{NNs+1}(n) &= U'_{NNs}(n); \quad \text{if } i = NTN, \text{ then } U_{NNTN+1}(n) = U_{Nr}(n) \\ F_{sin}(n) &= Q_{L_{sin}}(n-1) + Q'_{L_{NNs}}(n-1) + b'_{NNs}(n)V_{NNs}(n-1) - V_{C_{sin}}(n-1) \\ &= -b_{NNs}(n)U_{NNs}(n) + H_{sf}(n)U'_{NNs}(n) - b_{sin}U_{NNs+1}(n) \end{aligned}$$

$$\begin{aligned}
F_{NNs+1}(n) &= Q_{L_{NNs+1}}(n-1) - Q_{L_{sin}}(n-1) + V_{C_{sin}}(n-1) - b_{NNs+1}(n)V_{NNs+1}(n-1) \\
&= -b_{sin}U'_{NNs}(n) + H_{sb}(n)U_{NNs+1}(n) - b_{NNs+1}(n)U_{NNs+2}(n) \\
F_{Nr}(n) &= b_{NNTN}(n)V_{NNTN}(n-1) + b_{Nr}(n)V_{L_{Nr}}(n-1) + Q'_{L_{NNTN}}(n-1) \\
&= -b_{NNTN}(n)U_{NNTN}(n) + H_{Nr}(n)U_{Nr}(n) \\
P_{Ni}(n) &= b_{Ni}(n)[U_{Ni}(n) - U_{Ni+1}(n) + V_{Ni}(n-1)],
\end{aligned}$$

where

$$i = 1, 2, \ldots, NTN; \quad \text{if } i = Ns, \text{ then } U_{NNs+1}(n) = U'_{NNs}(n)$$

if

$$\begin{aligned}
&i = NTN, \quad \text{then } U_{NNTN+1}(n) = U_{Nr}(n) \\
P_{sin}(n) &= b_{sin}\left[U'_{NNs}(n) - U_{NNs+1}(n)\right] - Q_{L_{sin}}(n-1) + V_{C_{sin}}(n-1) \\
P_{Nr}(n) &= b_{Nr}(n)[U_{Nr}(n) - V_{L_{Nr}}(n-1)]
\end{aligned}$$

$$\begin{bmatrix} F_{N1} \\ F_{N2} \\ F_{N3} \\ \vdots \\ F_{sin} \\ F_{NNs+1} \\ \vdots \\ F_{NNTN} \\ F_{Nr} \end{bmatrix} = \begin{bmatrix} -1 & H_{N1} & -b_{N1} & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -b_{N1} & H_{N2} & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -b_{N2} & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -b_{NNs} & H_{sf} & -b_{sin} & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -b_{sin} & H_{sb} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & H_{NNTN} & -b_{NNTN} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & -b_{NNTN} & H_{Nr} \end{bmatrix} \begin{bmatrix} P_{NC} \\ U_{N1} \\ U_{N2} \\ \vdots \\ U'_{NNs} \\ U_{NNs+1} \\ \vdots \\ U_{NNTN} \\ U_{Nr} \end{bmatrix}$$

(A11–C.24)

If there is no nasal sinus cavity, the matrix Equation (A11–C.24) reduces to the following.

$$\begin{aligned}
F_{N1}(n) &= Q_{L_{N1}}(n-1) - b_{N1}(n)V_{N1}(n-1) \\
&= -P_{NC}(n) + H_{N1}(n)U_{N1}(n) - b_{N1}(n)U_{N2}(n) \\
F_{Ni}(n) &= Q_{L_{Ni-1}L_{Ni}}(n-1) + b_{Ni-1}(n)V_{Ni-1}(n-1) - b_{Ni}(n)V_{Ni}(n-1) \\
&= -b_{Ni-1}(n)U_{Ni-1}(n) + H_{Ni}(n)U_{Ni}(n) - b_{Ni}(n)U_{Ni+1}(n)
\end{aligned}$$

where

$$\begin{aligned}
&i = 2, \ldots, NTN; \quad \text{if } i = NTN, \text{ then } U_{NNTN+1}(n) = U_{Nr}(n) \\
F_{Nr}(n) &= b_{NNTN}(n)V_{NNTN}(n-1) + b_{Nr}(n)V_{L_{Nr}}(n-1) + Q'_{L_{NNTN}}(n-1) \\
&= -b_{NNTN}(n)U_{NNTN}(n) + H_{Nr}(n)U_{Nr}(n) \\
P_{Ni}(n) &= b_{Ni}(n)[U_{Ni}(n) - U_{Ni+1}(n) + V_{Ni}(n-1)],
\end{aligned}$$

where

$$\begin{aligned}
&i = 1, \ldots, NTN; \quad \text{if } i = NTN, \text{ then } U_{NNTN+1}(n) = U_{Nr}(n) \\
P_{Nr}(n) &= b_{Nr}(n)[U_{Nr}(n) - V_{L_{Nr}}(n-1)]
\end{aligned}$$

$$\begin{bmatrix} F_{N1} \\ F_{N2} \\ F_{N3} \\ \cdot \\ \cdot \\ \cdot \\ F_{NNTN} \\ F_{Nr} \end{bmatrix} = \begin{bmatrix} -1 & H_{N1} & -b_{N1} & 0 & \cdots & 0 & 0 & 0 \\ 0 & -b_{N1} & H_{N2} & -b_{N2} & \cdots & 0 & 0 & 0 \\ 0 & 0 & -b_{N2} & H_{N3} & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & -b_{NNTN-1} & H_{NNTN} & -b_{NNTN} \\ 0 & 0 & 0 & 0 & \cdots & 0 & -b_{NNTN} & H_{Nr} \end{bmatrix} \begin{bmatrix} P_{NC} \\ U_{N1} \\ U_{N2} \\ \cdot \\ \cdot \\ \cdot \\ U_{NNTN} \\ U_{Nr} \end{bmatrix}$$

(A11–C.25)

If the glottal impedance and the subglottal system are included, the matrix Equation (A11–C.22) extends to the following.

$$\begin{bmatrix} F_s \\ F_1 \\ F_2 \\ \cdot \\ \cdot \\ \cdot \\ F_{NTS} \\ F_{NC} \end{bmatrix} = \begin{bmatrix} 1 & X_{sub} & 0 & \cdots & 0 & 0 & 0 \\ -1 & H_1 & -b_1 & \cdots & 0 & 0 & 0 \\ 0 & -b_1 & H_2 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & H_{NTS} & -b_{NTS} & 0 \\ 0 & 0 & 0 & \cdots & -b_{NTS} & H_{NC} & 1 \end{bmatrix} \begin{bmatrix} P_g \\ U_1 \\ U_2 \\ \cdot \\ \cdot \\ \cdot \\ U_{NC} \\ P_{NC} \end{bmatrix} \quad \text{(A11–C.26)}$$

From the derived discrete-time acoustic matrix equations, we can form four different vocal system model structures, which are:

1. (A11–C.22), (A11–C.23), and (A11–C.24) for the vocal system model with nasal sinus but no glottal impedance and no subglottal system.
2. (A11–C.22), (A11–C.23), and (A11–C.25) for the vocal system model with no nasal sinus and no glottal impedance and no subglottal system.
3. (A11–C.26), (A11–C.23), and (A11–C.24) for the vocal system model with nasal sinus and with the glottal impedance and the subglottal system.
4. (A11–C.26), (A11–C.23), and (A11–C.25) for the vocal system model with no nasal sinus but with the glottal impedance and the subglottal system.

There are three matrix equations for each structure. Each matrix equation can be written as $y = A \cdot x$, where A is a non-square band diagonal sparse matrix of coefficients, y is a column vector of force constants, and x is the unknown column vector. The elements in the sparse coefficient matrix and the force constants inside the vocal system are defined in the previous paragraphs. In addition to the three matrix equations, we need the boundary condition at the nasal coupling point, $U_{NC}(n) = U_{i+1}(n) + U_{N1}(n)$. For Cases 3 and 4, we need one more boundary condition at the glottis, $U_g(n) = U_1(n) + U_0(n)$. It may also be noted that the three matrix equations for each case are coupled by the term $P_{NC}(n)$ and boundary condition(s). As Maeda (1982a) pointed out, if we eliminate $P_{NC}(n)$ and $U_{NC}(n)$ analytically, the formulation results in an unstable system. Therefore, $P_{NC}(n)$ and $U_{NC}(n)$ are included in the unknown column vectors in order for stable solutions.

There are many methods that can be used to solve these sparse linear system equations. Since the coefficient matrices are band diagonal sparse matrices, an efficient elimination procedure followed by a substitution procedure can be used. Once the values of $U_i(n)$ and $U_{Ni}(n)$ for all i, $P_{NC}(n)$, and $U_{NC}(n)$ are solved, the pressure $P_i(n)$ and $P_{Ni}(n)$ for all i, and the volume velocity $u_{i3}(n)$ and $u_{Ni3}(n)$ for all i can be computed. Then, the force constants and coefficient matrices are updated for the next recursion.

APPENDIX **11–D**

COMMENTS ON USAGE OF SIMULATED ANNEALING PROCEDURE

In Appendix 11, we covered the details of speech inverse filtering using the simulated annealing optimization algorithm. The annealing parameters control the performance of the optimization process. The various acoustic characteristics of the speech signal and target-frame selections affect the optimization process. The default values of the annealing parameters are not the best values for all cases. Although we do not know the combinations of the annealing parameters that perform well for the optimization process for different target frames, the following guidelines provide some rules for adjusting the appropriate annealing parameters.

Step 1: Set the desired nasalization extent and set the number of dimensions of the articulatory vector at the appropriate dimensions, for example, M = 8 for front vowels; M = 9 for nasalized front vowels; M = 11 for middle, back vowels, and semivowels; and M = 12 for nasalized vowels (see the descriptions in Section A11.3.3.1). Start the optimization process with the default initial articulatory vector and the default annealing parameters. If the error distance is less than 1% after the process stops, go to Step 5. If not, go to Step 2.

Step 2: Check if the current vocal tract shape (or cross-sectional area) is reasonable. If the shape is not reasonable go to Step 3. Otherwise, record the error distance as ε_p and the current final temperature as T_p. Then set the initial temperature $T = \text{floor}[T_p]$. Start the optimization process again. If the new error distance is less than ε_p, then this step is repeated until the error criterion is met. If not go to Step 4.

Step 3: Examine the vocal tract shape (articulatory vector) for the initial settings and all subsequent values. Several adjustments of the annealing parameters can be used. The following order of adjustments are recommended: raise the initial temperature, increase the value of the reduction factor, increase the total number of evaluations, and change the other annealing parameters as desired. Then begin the process and apply Step 2.

Step 4: Examine the vocal tract shape as described in Step 3. Increase the number of dimensions of the articulatory vector from M = 8 to M = 11 and start the process. Apply Step 2.

Step 5: Check the vocal tract shape with x-ray tracings or schematic vocal tract profiles as published in the literature or compare with vocal tract profiles as shown in Appendix 5. If the vocal tract outline is similar to those published, then the optimization process is done. If not, this may mean that the "ventriloquist effect" has occurred. One can adjust the settings of the nine articulatory sliders in the Articulatory Position Settings popup window so that the initial configuration becomes closer to the true outline. Then go to Step 1 to start the optimization process again.

Step 6: If the above steps fail to reduce the error criterion to a satisfactory value, then return to the target-frame selection phase and reselect the current target frame. Start the optimization process from the Step 1.

The above guidelines are illustrated in the following examples for the sentence, "We were away a year ago," spoken by two male subjects, A and B. Figures A11–D.1 and A11–D.2 show the optimized target frames for subjects A and B, respectively. See Appendix A11–A for the vowel results.

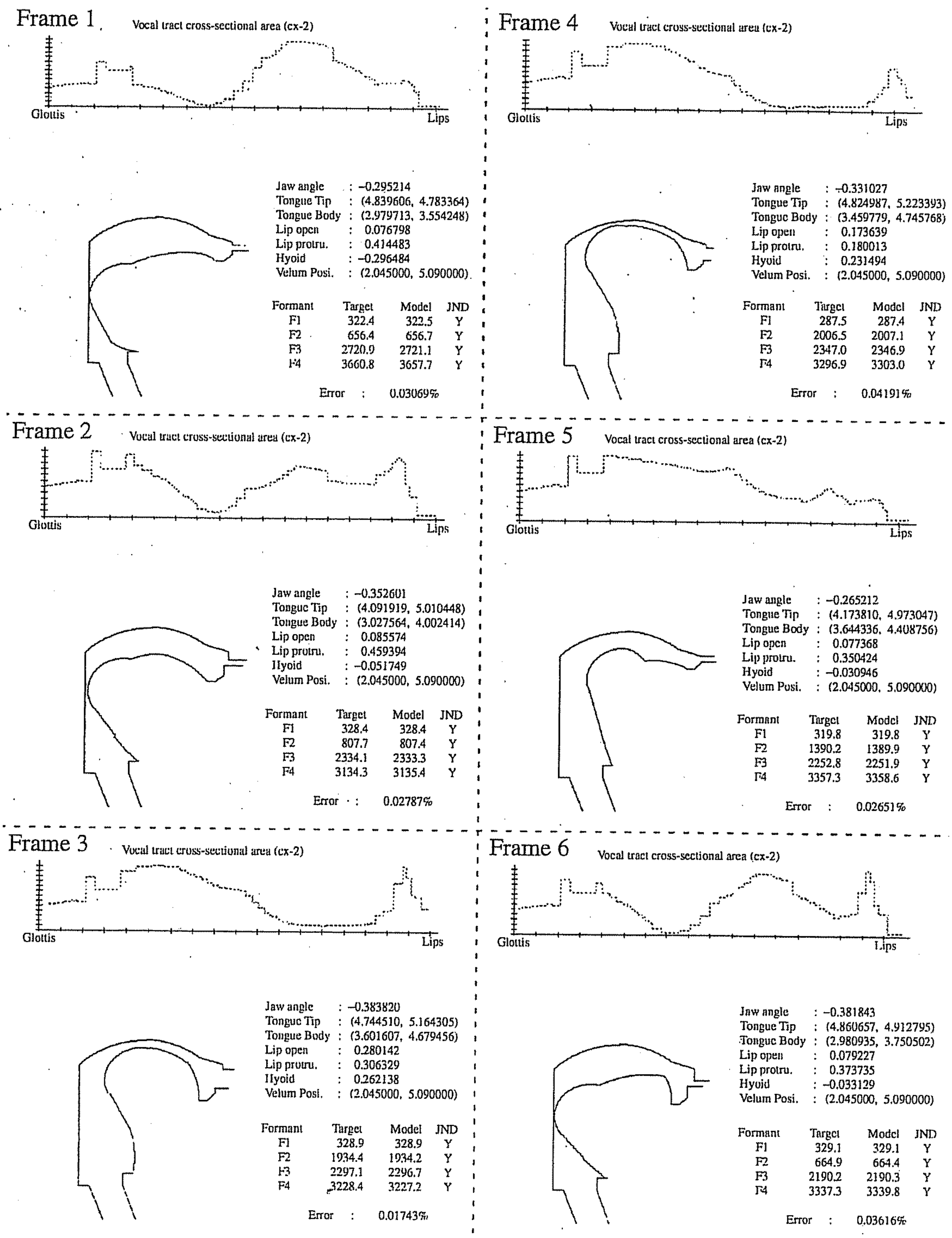
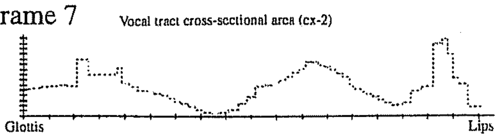
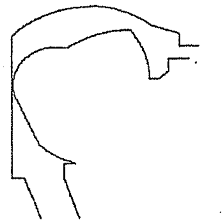


FIGURE A11-D.1 The optimized target frames of the sentence, "We were away a year ago," spoken by male subject A.

Frame 7 Vocal tract cross-sectional area (cx-2)



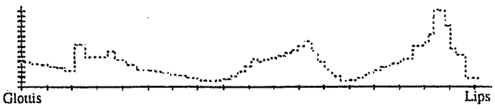


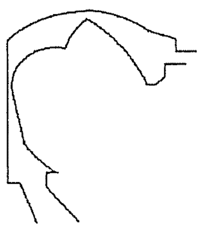
Jaw angle : –0.400354
Tongue Tip : (4.581029, 5.228440)
Tongue Body : (2.984698, 3.880309)
Lip open : 0.158502
Lip protru. : 0.419453
Hyoid : –0.293551
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	383.6	383.8	Y
F2	813.8	813.3	Y
F3	1890.2	1889.5	Y
F4	3259.6	3264.4	Y

Error : 0.06137%

Frame 8 Vocal tract cross-sectional area (cx-2)



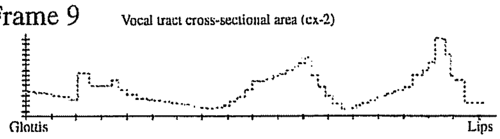


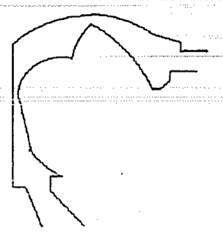
Jaw angle : –0.401027
Tongue Tip : (3.710760, 5.614390)
Tongue Body : (3.044841, 4.003826)
Lip open : 0.159859
Lip protru. : 0.428789
Hyoid : –0.285939
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	423.7	423.7	Y
F2	1003.4	1003.5	Y
F3	1617.6	1617.7	Y
F4	3195.1	3196.1	Y

Error : 0.00955%

Frame 9 Vocal tract cross-sectional area (cx-2)



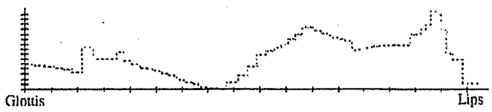


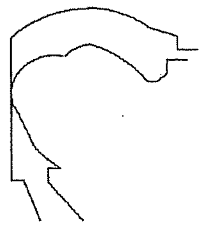
Jaw angle : –0.401388
Tongue Tip : (3.700464, 5.577478)
Tongue Body : (3.079889, 3.860703)
Lip open : 0.239711
Lip protru. : 0.564889
Hyoid : –0.298830
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	450.7	450.0	Y
F2	1064.7	1064.8	Y
F3	1688.0	1688.3	Y
F4	3285.1	3286.1	Y

Error : 0.05414%

Frame 10 Vocal tract cross-sectional area (cx-2)



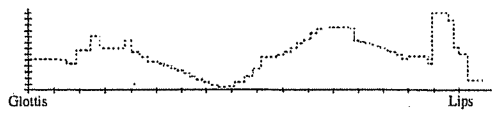


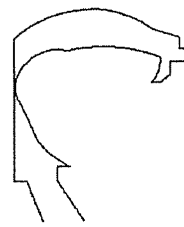
Jaw angle : –0.401008
Tongue Tip : (3.736150, 4.931845)
Tongue Body : (2.958587, 3.740478)
Lip open : 0.129357
Lip protru. : 0.429982
Hyoid : –0.168980
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	376.2	376.2	Y
F2	631.4	631.6	Y
F3	2238.1	2238.3	Y
F4	3182.7	3298.0	Y

Error : 0.55861%

Frame 11 Vocal tract cross-sectional area (cx-2)



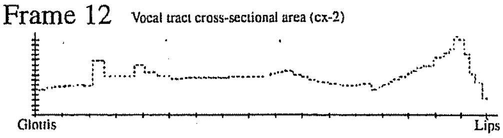


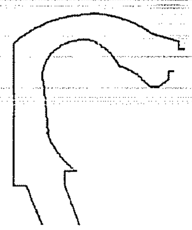
Jaw angle : –0.396526
Tongue Tip : (5.208612, 4.633890)
Tongue Body : (2.990479, 3.917469)
Lip open : 0.152407
Lip protru. : 0.383981
Hyoid : –0.205112
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	413.3	413.3	Y
F2	815.4	815.6	Y
F3	2312.4	2312.2	Y
F4	3186.7	3187.7	Y

Error : 0.01406%

Frame 12 Vocal tract cross-sectional area (cx-2)





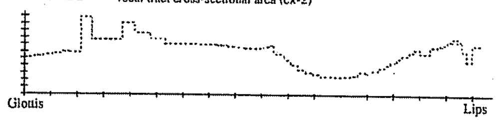
Jaw angle : –0.400731
Tongue Tip : (4.331834, 4.522106)
Tongue Body : (3.575943, 4.261568)
Lip open : 0.266536
Lip protru. : 0.113469
Hyoid : –0.286778
Velum Posi. : (2.045000, 5.090000)

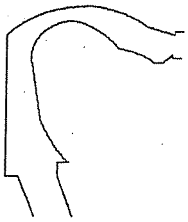
Formant	Target	Model	JND
F1	500.3	500.1	Y
F2	1411.4	1411.6	Y
F3	2203.5	2204.7	Y
F4	3354.0	3355.4	Y

Error : 0.03203%

FIGURE A11–D.1 Continued.

Frame 13 Vocal tract cross-sectional area (cx-2)



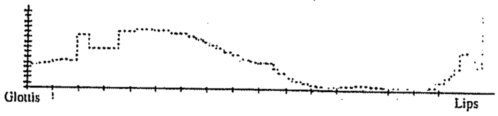


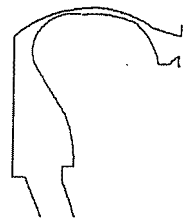
Jaw angle	:	–0.320191
Tongue Tip	:	(4.438051, 4.767700)
Tongue Body	:	(3.474534, 4.529708)
Lip open	:	0.330985
Lip protru.	:	0.195973
Hyoid	:	–0.247216
Velum Posi.	:	(2.045000, 5.090000)

Formant	Target	Model	JND
F1	450.0	450.0	Y
F2	1713.9	1713.5	Y
F3	2339.7	2340.4	Y
F4	3479.3	3479.2	Y

Error : 0.01556%

Frame 16 Vocal tract cross-sectional area (cx-2)



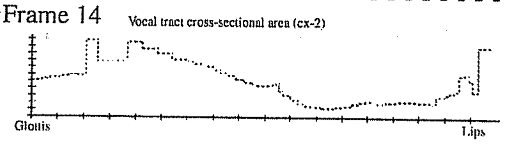


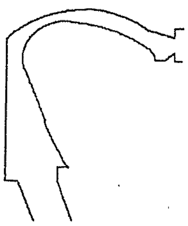
Jaw angle	:	–0.318895
Tongue Tip	:	(4.703774, 5.410002)
Tongue Body	:	(3.345196, 4.732948)
Lip open	:	0.498416
Lip protru.	:	0.038157
Hyoid	:	0.202666
Velum Posi.	:	(2.045000, 5.090000)

Formant	Target	Model	JND
F1	275.2	275.3	Y
F2	2086.6	2086.5	Y
F3	2871.7	2871.6	Y
F4	3699.3	3699.2	Y

Error : 0.00472%

Frame 14 Vocal tract cross-sectional area (cx-2)



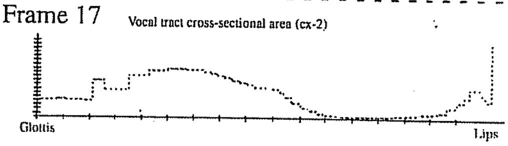


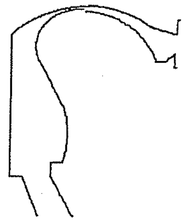
Jaw angle	:	–0.296191
Tongue Tip	:	(5.070441, 4.866640)
Tongue Body	:	(3.302177, 4.607279)
Lip open	:	0.408096
Lip protru.	:	0.178389
Hyoid	:	–0.065855
Velum Posi.	:	(2.045000, 5.090000)

Formant	Target	Model	JND
F1	370.7	371.0	Y
F2	1887.4	1888.7	Y
F3	2539.3	2539.7	Y
F4	3577.7	3576.0	Y

Error : 0.05564%

Frame 17 Vocal tract cross-sectional area (cx-2)



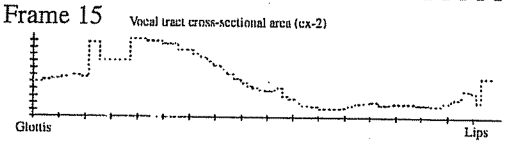


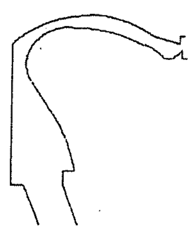
Jaw angle	:	–0.316181
Tongue Tip	:	(4.556979, 5.360120)
Tongue Body	:	(3.480864, 4.809723)
Lip open	:	0.575142
Lip protru.	:	0.045386
Hyoid	:	0.295078
Velum Posi.	:	(2.045000, 5.090000)

Formant	Target	Model	JND
F1	274.5	274.5	Y
F2	2128.8	2128.8	Y
F3	3281.7	3281.4	Y
F4	3612.5	3613.4	Y

Error : 0.00951%

Frame 15 Vocal tract cross-sectional area (cx-2)



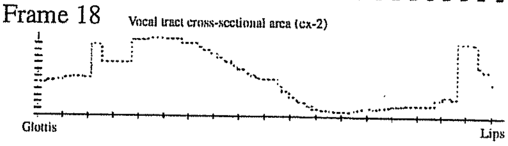


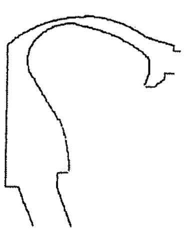
Jaw angle	:	–0.263461
Tongue Tip	:	(4.864736, 5.085028)
Tongue Body	:	(3.253271, 4.595029)
Lip open	:	0.274938
Lip protru.	:	0.105571
Hyoid	:	0.127595
Velum Posi.	:	(2.045000, 5.090000)

Formant	Target	Model	JND
F1	346.1	346.2	Y
F2	1873.4	1874.3	Y
F3	2511.8	2512.0	Y
F4	3565.4	3564.1	Y

Error : 0.03051%

Frame 18 Vocal tract cross-sectional area (cx-2)





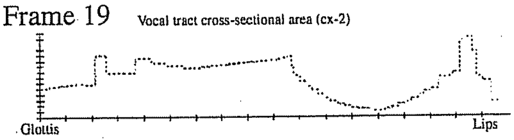
Jaw angle	:	–0.381602
Tongue Tip	:	(5.112584, 4.829454)
Tongue Body	:	(3.391806, 4.715267)
Lip open	:	0.277634
Lip protru.	:	0.167777
Hyoid	:	0.197406
Velum Posi.	:	(2.045000, 5.090000)

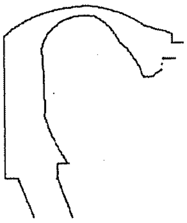
Formant	Target	Model	JND
F1	317.3	317.3	Y
F2	2001.0	2000.6	Y
F3	2336.6	2336.8	Y
F4	3343.2	3343.5	Y

Error : 0.01158%

FIGURE A11–D.1 Continued.

Frame 19 Vocal tract cross-sectional area (cx-2)



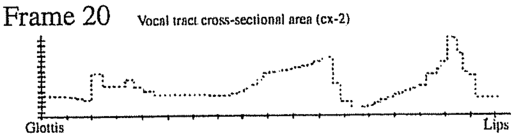


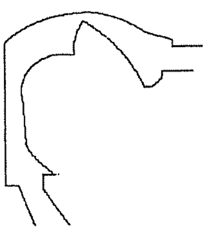
Jaw angle : −0.390341
Tongue Tip : (4.500666, 4.943980)
Tongue Body : (3.747747, 4.680444)
Lip open : 0.189774
Lip protru. : 0.251803
Hyoid : −0.241062
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	365.1	365.2	Y
F2	1562.8	1562.7	Y
F3	1889.9	1890.7	Y
F4	3371.3	3370.0	Y

Error : 0.02512%

Frame 20 Vocal tract cross-sectional area (cx-2)



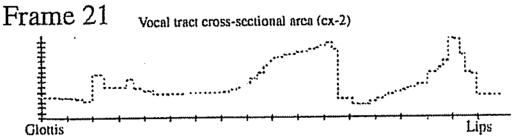


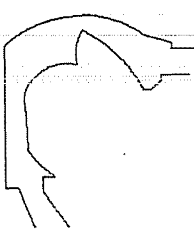
Jaw angle : −0.401420
Tongue Tip : (3.700176, 5.597184)
Tongue Body : (3.293831, 3.886328)
Lip open : 0.297621
Lip protru. : 0.647351
Hyoid : −0.299500
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	445.8	445.4	Y
F2	1277.3	1276.9	Y
F3	1710.9	1722.5	Y
F4	3261.8	3261.0	Y

Error : 0.21271%

Frame 21 Vocal tract cross-sectional area (cx-2)



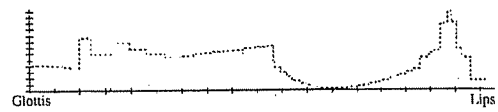


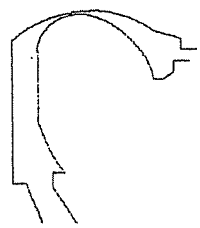
Jaw angle : −0.401201
Tongue Tip : (3.700115, 5.443300)
Tongue Body : (3.376132, 3.720156)
Lip open : 0.321179
Lip protru. : 0.609505
Hyoid : −0.298412
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	478.6	478.6	Y
F2	1323.2	1323.0	Y
F3	1910.2	1910.0	Y
F4	3311.2	3311.1	Y

Error : 0.00853%

Frame 22 Vocal tract cross-sectional area (cx-2)



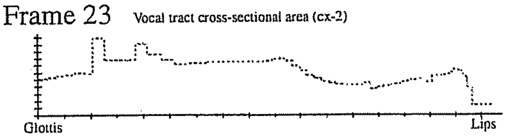


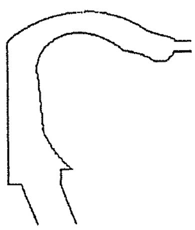
Jaw angle : −0.378224
Tongue Tip : (4.562151, 5.107852)
Tongue Body : (3.487845, 4.818657)
Lip open : 0.147268
Lip protru. : 0.352104
Hyoid : −0.299674
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	300.5	300.5	Y
F2	1295.4	1295.5	Y
F3	1827.1	1827.2	Y
F4	3807.5	3565.1	Y

Error : 0.95769%

Frame 23 Vocal tract cross-sectional area (cx-2)



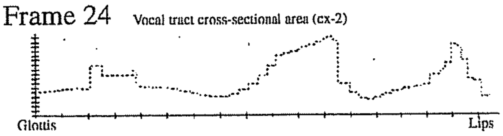


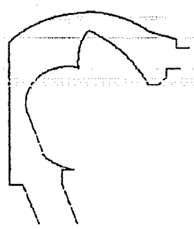
Jaw angle : −0.297112
Tongue Tip : (4.524280, 4.859411)
Tongue Body : (3.563238, 4.394766)
Lip open : 0.125476
Lip protru. : 0.343277
Hyoid : −0.284596
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	387.7	387.9	Y
F2	1373.2	1373.2	Y
F3	2176.9	2177.1	Y
F4	3363.7	3364.0	Y

Error : 0.01419%

Frame 24 Vocal tract cross-sectional area (cx-2)





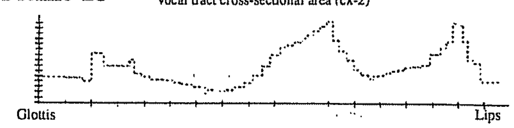
Jaw angle : −0.400097
Tongue Tip : (3.708182, 5.350190)
Tongue Body : (3.309445, 3.573624)
Lip open : 0.253914
Lip protru. : 0.275782
Hyoid : −0.295448
Velum Posi. : (2.045000, 5.090000)

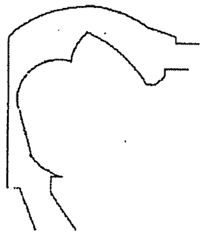
Formant	Target	Model	JND
F1	480.2	480.2	Y
F2	1238.7	1238.7	Y
F3	2149.4	2149.1	Y
F4	3365.5	3365.2	Y

Error : 0.00563%

FIGURE A11-D.1 Continued.

Frame 25 Vocal tract cross-sectional area (cx-2)



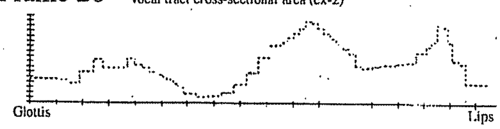


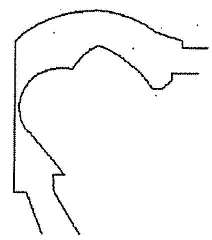
Jaw angle : -0.400450
Tongue Tip : (3.701344, 5.207943)
Tongue Body : (3.157058, 3.640734)
Lip open : 0.296594
Lip protru. : 0.497078
Hyoid : -0.299717
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	518.4	518.4	Y
F2	1126.7	1129.3	Y
F3	2196.0	2196.2	Y
F4	3398.9	3401.4	Y

Error : 0.08341%

Frame 26 Vocal tract cross-sectional area (cx-2)





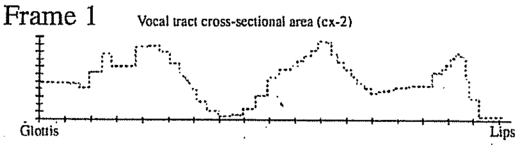
Jaw angle : -0.400763
Tongue Tip : (3.858089, 4.994038)
Tongue Body : (3.056451, 3.561519)
Lip open : 0.294234
Lip protru. : 0.547207
Hyoid : 0.012368
Velum Posi. : (2.045000, 5.090000)

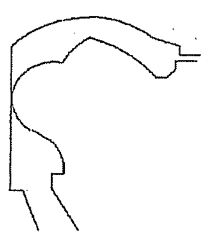
Formant	Target	Model	JND
F1	504.5	504.3	Y
F2	987.8	987.8	Y
F3	2329.8	2326.9	Y
F4	3416.9	3417.4	Y

Error : 0.04494%

FIGURE A11-D.1 Continued.

Frame 1 Vocal tract cross-sectional area (cx-2)



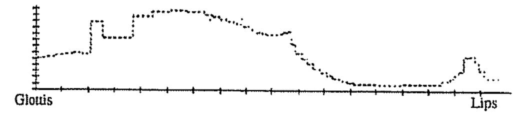


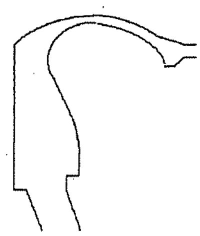
Jaw angle : -0.335120
Tongue Tip : (3.703419, 5.311586)
Tongue Body : (2.992716, 3.823177)
Lip open : 0.070736
Lip protru. : 0.451054
Hyoid : 0.247373
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	308.2	308.2	Y
F2	725.6	725.6	Y
F3	1973.4	1973.3	Y
F4	3134.4	3131.1	Y

Error : 0.02402%

Frame 3 Vocal tract cross-sectional area (cx-2)



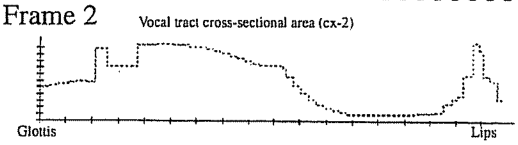


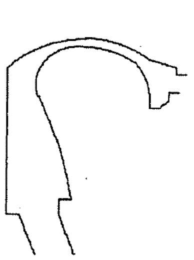
Jaw angle : -0.289567
Tongue Tip : (4.989875, 5.056913)
Tongue Body : (3.659096, 4.719562)
Lip open : 0.153200
Lip protru. : 0.257071
Hyoid : 0.195418
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	297.4	297.4	Y
F2	1894.2	1894.4	Y
F3	2338.4	2338.7	Y
F4	3278.2	3272.6	Y

Error : 0.03528%

Frame 2 Vocal tract cross-sectional area (cx-2)



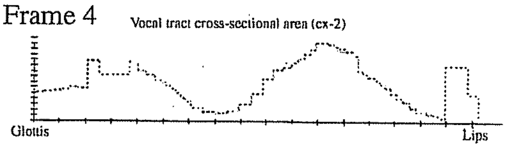


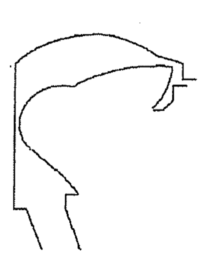
Jaw angle : -0.373695
Tongue Tip : (4.811612, 5.183755)
Tongue Body : (3.576164, 4.719492)
Lip open : 0.212940
Lip protru. : 0.197551
Hyoid : 0.094424
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	311.0	310.9	Y
F2	1916.5	1917.2	Y
F3	2258.7	2259.0	Y
F4	3282.3	3283.0	Y

Error : 0.02022%

Frame 4 Vocal tract cross-sectional area (cx-2)





Jaw angle : -0.397331
Tongue Tip : (5.414547, 5.054583)
Tongue Body : (3.069459, 3.678089)
Lip open : 0.063201
Lip protru. : 0.306152
Hyoid : 0.036285
Velum Posi. : (2.045000, 5.090000)

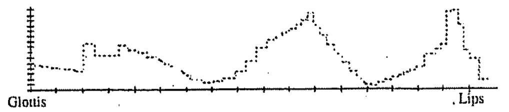
Formant	Target	Model	JND
F1	309.5	309.5	Y
F2	852.5	852.5	Y
F3	1789.8	1790.3	Y
F4	3300.0	3299.2	Y

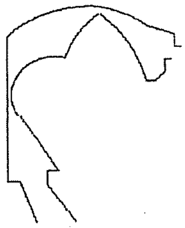
Error : 0.01403%

FIGURE A11-D.2 The optimized target frames of the sentence, "We were away a year ago," spoken by male subject B.

Frame 5

Vocal tract cross-sectional area (cx-2)





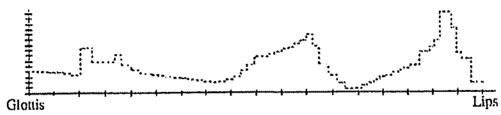
Jaw angle : −0.398308
Tongue Tip : (4.012872, 5.636487)
Tongue Body : (3.060456, 3.686242)
Lip open : 0.142035
Lip protru. : 0.144014
Hyoid : 0.019017
Velum Posi. : (2.045000, 5.090000)

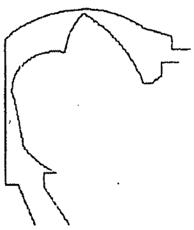
Formant	Target	Model	JND
F1	388.8	388.7	Y
F2	1011.7	1011.6	Y
F3	1574.5	1573.6	Y
F4	3314.2	3316.9	Y

Error : 0.03175%

Frame 6

Vocal tract cross-sectional area (cx-2)





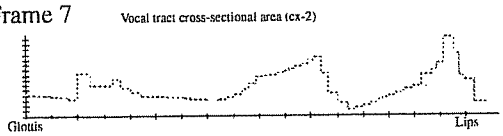
Jaw angle : −0.400235
Tongue Tip : (3.712281, 5.706581)
Tongue Body : (3.106781, 3.889300)
Lip open : 0.153682
Lip protru. : 0.386648
Hyoid : −0.299157
Velum Posi. : (2.045000, 5.090000)

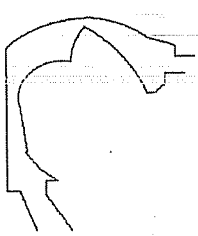
Formant	Target	Model	JND
F1	397.0	397.0	Y
F2	1084.3	1083.8	Y
F3	1440.6	1441.2	Y
F4	3330.0	3332.6	Y

Error : 0.03570%

Frame 7

Vocal tract cross-sectional area (cx-2)





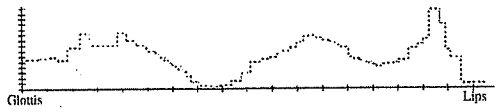
Jaw angle : −0.401159
Tongue Tip : (3.701498, 5.597988)
Tongue Body : (3.202925, 3.875268)
Lip open : 0.204319
Lip protru. : 0.409777
Hyoid : −0.297142
Velum Posi. : (2.045000, 5.090000)

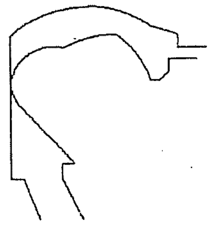
Formant	Target	Model	JND
F1	452.7	448.6	Y
F2	1207.1	1207.1	Y
F3	1706.7	1713.4	Y
F4	3281.2	3289.4	Y

Error : 0.41168%

Frame 8

Vocal tract cross-sectional area (cx-2)





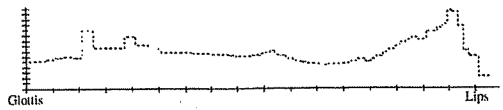
Jaw angle : −0.397946
Tongue Tip : (4.259106, 5.130637)
Tongue Body : (2.977700, 3.916171)
Lip open : 0.130780
Lip protru. : 0.588537
Hyoid : 0.020919
Velum Posi. : (2.045000, 5.090000)

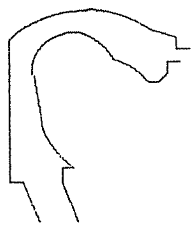
Formant	Target	Model	JND
F1	329.8	329.9	Y
F2	680.4	680.4	Y
F3	1965.9	1965.7	Y
F4	3114.7	3119.6	Y

Error : 0.03730%

Frame 9

Vocal tract cross-sectional area (cx-2)





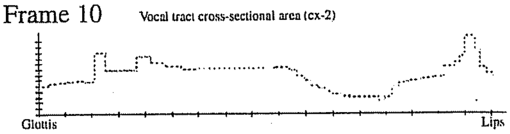
Jaw angle : −0.392355
Tongue Tip : (4.247635, 4.623539)
Tongue Body : (3.442429, 4.342550)
Lip open : 0.164341
Lip protru. : 0.264943
Hyoid : −0.276863
Velum Posi. : (2.045000, 5.090000)

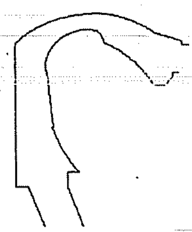
Formant	Target	Model	JND
F1	453.6	453.4	Y
F2	1239.5	1239.4	Y
F3	2142.4	2143.0	Y
F4	3341.7	3341.3	Y

Error : 0.02390%

Frame 10

Vocal tract cross-sectional area (cx-2)





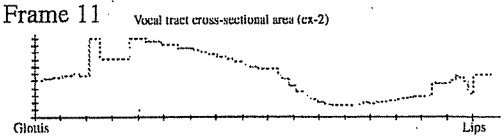
Jaw angle : −0.389529
Tongue Tip : (3.966737, 5.098867)
Tongue Body : (3.608622, 4.517625)
Lip open : 0.335497
Lip protru : 0.135433
Hyoid : −0.223796
Velum Posi : (2.045000, 5.090000)

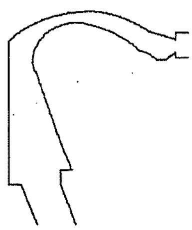
Formant	Target	Model	JND
F1	464.0	464.0	Y
F2	1704.9	1705.3	Y
F3	2267.2	2267.2	Y
F4	3423.2	3423.7	Y

Error : 0.01066%

FIGURE A11–D.2 Continued.

Frame 11 Vocal tract cross-sectional area (cx-2)



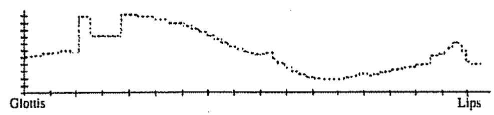


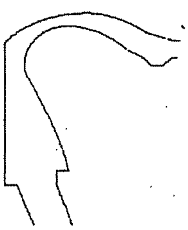
Jaw angle : −0.288207
Tongue Tip : (4.865392, 4.844986)
Tongue Body : (3.478848, 4.624095)
Lip open : 0.306237
Lip protru. : 0.268265
Hyoid : −0.011991
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	376.7	376.7	Y
F2	1856.3	1856.0	Y
F3	2417.9	2418.2	Y
F4	3423.7	3424.4	Y

Error : 0.01196%

Frame 12 Vocal tract cross-sectional area (cx-2)



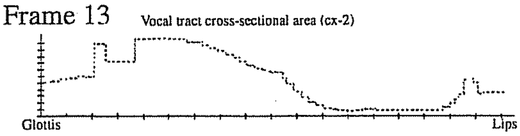


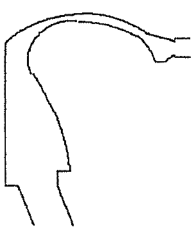
Jaw angle : −0.309652
Tongue Tip : (4.659431, 4.892226)
Tongue Body : (3.386737, 4.583634)
Lip open : 0.219131
Lip protru. : 0.129724
Hyoid : 0.077292
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	388.4	388.2	Y
F2	1798.2	1798.3	Y
F3	2345.4	2345.4	Y
F4	3396.4	3395.8	Y

Error : 0.01656%

Frame 13 Vocal tract cross-sectional area (cx-2)



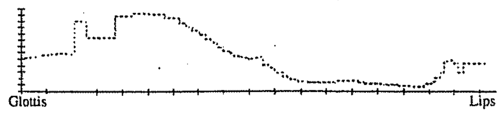


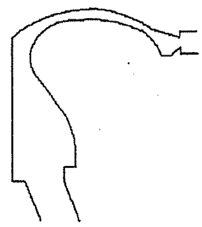
Jaw angle : −0.288301
Tongue Tip : (4.823072, 5.259611)
Tongue Body : (3.461715, 4.731454)
Lip open : 0.232324
Lip protru. : 0.340106
Hyoid : 0.148036
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	288.3	288.3	Y
F2	1981.7	1982.1	Y
F3	2513.4	2513.4	Y
F4	3330.0	3331.6	Y

Error : 0.01827%

Frame 14 Vocal tract cross-sectional area (cx-2)



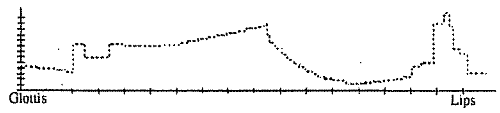


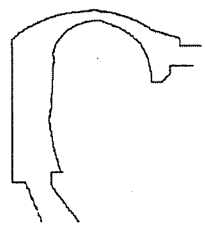
Jaw angle : −0.277270
Tongue Tip : (4.822687, 5.273158)
Tongue Body : (3.353062, 4.633030)
Lip open : 0.272794
Lip protru. : 0.355602
Hyoid : 0.240086
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	307.6	307.6	Y
F2	1916.4	1917.0	Y
F3	2558.2	2557.8	Y
F4	3360.8	3359.0	Y

Error : 0.02434%

Frame 15 Vocal tract cross-sectional area (cx-2)



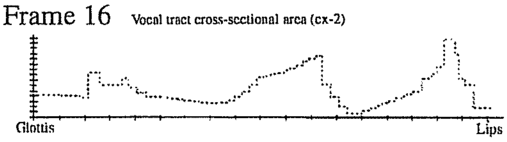


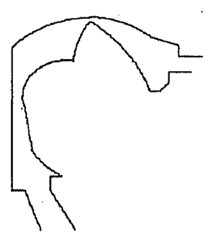
Jaw angle : −0.389651
Tongue Tip : (4.740730, 5.042890)
Tongue Body : (3.820071, 4.648295)
Lip open : 0.237129
Lip protru. : 0.454148
Hyoid : −0.181307
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	371.1	371.1	Y
F2	1647.3	1646.9	Y
F3	1924.3	1923.2	Y
F4	3308.6	3311.9	Y

Error : 0.03579%

Frame 16 Vocal tract cross-sectional area (cx-2)





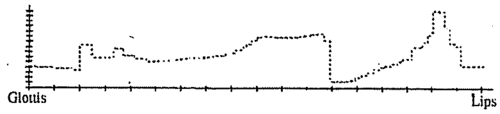
Jaw angle : −0.398010
Tongue Tip : (3.728233, 5.690279)
Tongue Body : (3.166779, 3.855053)
Lip open : 0.183063
Lip protru. : 0.491582
Hyoid : −0.294449
Velum Posi. : (2.045000, 5.090000)

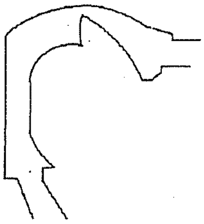
Formant	Target	Model	JND
F1	399.2	399.2	Y
F2	1133.0	1133.6	Y
F3	1494.4	1495.1	Y
F4	3305.8	3305.5	Y

Error : 0.02799%

FIGURE A11-D.2 Continued.

Frame 17 Vocal tract cross-sectional area (cx-2)



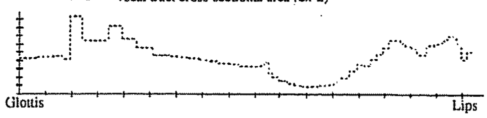


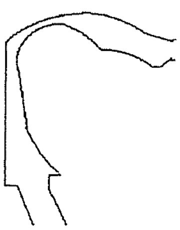
Jaw angle : −0.400712
Tongue Tip : (3.709026, 5.578534)
Tongue Body : (3.479430, 3.951865)
Lip open : 0.325082
Lip protru. : 0.595860
Hyoid : −0.283749
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	440.4	440.7	Y
F2	1415.7	1415.8	Y
F3	1832.3	1830.4	Y
F4	3270.3	3204.9	Y

Error : 0.35461%

Frame 18 Vocal tract cross-sectional area (cx-2)



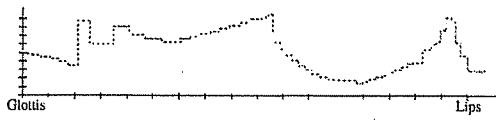


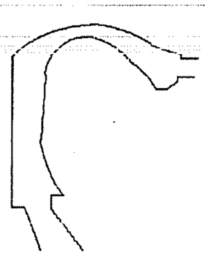
Jaw angle : −0.306939
Tongue Tip : (4.118007, 4.922364)
Tongue Body : (3.211027, 4.624109)
Lip open : 0.254570
Lip protru. : 0.091089
Hyoid : −0.297978
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	417.8	417.8	Y
F2	1643.1	1643.3	Y
F3	2186.1	2186.6	Y
F4	3724.0	3668.4	Y

Error : 0.23163%

Frame 19 Vocal tract cross-sectional area (cx-2)



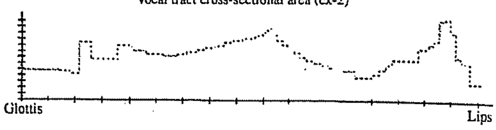


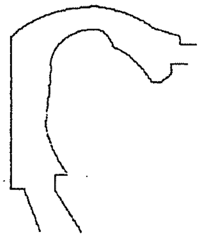
Jaw angle : −0.339906
Tongue Tip : (4.557400, 4.988651)
Tongue Body : (3.649744, 4.632542)
Lip open : 0.213379
Lip protru. : 0.367967
Hyoid : −0.286820
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	402.5	402.5	Y
F2	1608.0	1608.0	Y
F3	2006.4	2008.4	Y
F4	3364.3	3364.2	Y

Error : 0.02850%

Frame 20 Vocal tract cross-sectional area (cx-2)



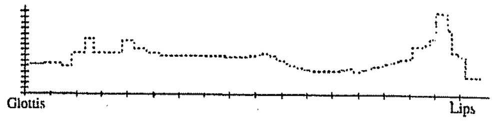


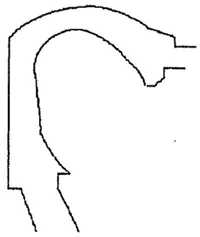
Jaw angle : −0.385660
Tongue Tip : (4.214625, 4.894730)
Tongue Body : (3.769741, 4.356261)
Lip open : 0.220392
Lip protru. : 0.347459
Hyoid : −0.279147
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	470.4	470.5	Y
F2	1486.7	1484.7	Y
F3	2110.8	2110.6	Y
F4	3348.5	3348.3	Y

Error : 0.00956%

Frame 21 Vocal tract cross-sectional area (cx-2)



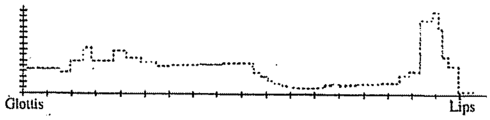


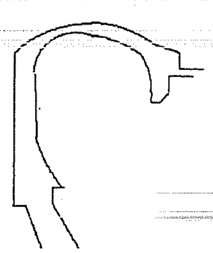
Jaw angle : −0.401105
Tongue Tip : (4.589791, 4.599903)
Tongue Body : (3.504574, 4.374523)
Lip open : 0.243942
Lip protru. : 0.425875
Hyoid : −0.265436
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	461.8	461.8	Y
F2	1383.6	1383.7	Y
F3	2106.4	2106.4	Y
F4	3263.8	3263.9	Y

Error : 0.00440%

Frame 22 Vocal tract cross-sectional area (cx-2)





Jaw angle : −0.401240
Tongue Tip : (4.822456, 5.040797)
Tongue Body : (3.387999, 4.662027)
Lip open : 0.088662
Lip protru. : 0.495322
Hyoid : −0.287198
Velum Posi. : (2.045000, 5.090000)

Formant	Target	Model	JND
F1	319.4	319.5	Y
F2	915.3	915.3	Y
F3	1878.8	1878.8	Y
F4	3156.5	3160.9	Y

Error : 0.03791%

FIGURE A11-D.2 Continued.