

# Effects of a curved vocal tract with grid-generated tongue profile on low-order formants

Paul H. Milenkovic<sup>a)</sup>

Department of Electrical and Computer Engineering, University of Wisconsin-Madison, 1415 Engineering Drive, Madison, Wisconsin 53706

Srikanth Yaddanapudi

CD-Adapco, No. 1222, 1801 Warner Ranch Drive, Round Rock, Texas 78664

Houri K. Vorperian

Waisman Center, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, Wisconsin 53705

Raymond D. Kent

Department of Communicative Disorders, University of Wisconsin-Madison, 1875 Willow Drive, Madison, Wisconsin 53706

(Received 4 February 2008; revised 22 October 2009; accepted 30 November 2009)

A hyperbolic grid-generation algorithm allows investigation of the effect of vocal-tract curvature on low-order formants. A smooth two-dimensional (2D) curve represents the combined lower lip, tongue, and anterior pharyngeal wall profile as displacements from the combined upper lip, palate, and posterior pharyngeal wall outline. The algorithm is able to generate tongue displacements beyond the local radius of strongly curved sections of the palate. The 2D grid, along with transverse profiles of the lip, oral-pharyngeal, and epilarynx regions, specifies a vocal conduit from which an effective area function may be determined using corrections to acoustic parameters resulting from duct curvature; the effective area function in turn determines formant frequencies through an acoustic transmission-line calculation. Results of the corrected transmission line are compared with a three-dimensional finite element model. The observed effects of the curved vocal tract on formants F1 and F2 are in order of importance, as follows: (1) reduction in midline distances owing to curvature of the palate and the bend joining the palate to the pharynx, (2) the curvature correction to areas and section lengths, and (3) adjustments to the palate-tongue distance required to produce smooth tongue shapes at large displacements from the palate.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3277214]

PACS number(s): 43.70.Aj, 43.70.Bk, 43.70.Dn [CHS]

Pages: 1002–1013

## I. INTRODUCTION

Recent studies<sup>1–4</sup> have investigated vocal-tract shape changes during child development using an articulatory model to compute the formant working space for different sizes and proportions of vocal tract. Acoustic resonator theory predicts that the formant working space scales linearly with the reciprocal of overall vocal tract length, and this is the dominant effect seen. The changes in vocal-tract proportions occurring during development are also hypothesized to affect the formant frequencies. The aim of the present study is to gain insights into the acoustic mechanisms by which variations in the outer vocal-tract outline apart from length changes could shift formant frequencies. To achieve this aim, the method of hyperbolic grid generation is adapted to determining the tongue shape in relation to a measured outline.

A common procedure in the child-development studies is to derive an outer vocal-tract outline, a rectilinear-radial grid pattern, and a set of tongue-displacement basis functions

from a reference adult subject. The outer vocal-tract outline refers to the contiguous outlines of upper lip, palate, and posterior pharyngeal wall in the midsagittal plane. These tongue-displacement basis functions can be derived from analysis of x-ray images<sup>5,6</sup> or be mathematically determined.<sup>7</sup> Manual identification of vocal-tract landmarks determines the grid pattern, and it is helpful to have a reference subject without a strongly curved palate outline so that the radial grid lines remain nearly normal to the palate and allow a sufficient range of displacement of the tongue before the radial lines converge. The oral and pharyngeal regions of this model may be differentially scaled to represent a child vocal tract at different stages of development by proportioning the rectilinear-radial grid pattern.

This procedure not only changes the overall vocal-tract length as well as the proportions between oral and pharyngeal regions but it also scales the basis functions controlling tongue shape in proportion to the scale changes in the respective vocal-tract regions. Boe *et al.*<sup>1</sup> ascribed the acoustic effect of changing vocal-tract proportions primarily to the scaling of the basis functions. The goal of the current investigation is to determine changes in formant frequency intrinsic

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: phmilenk@wisc.edu

sis to changes in the curvature and proportions of the outer vocal-tract outline apart from basis function scaling and other influences.

An investigation into mechanisms by which vocal-tract proportions affect the formant frequencies depends on a model for the acoustic properties of a curved vocal tract. Morse and Ingard<sup>8</sup> along with Sundberg *et al.*<sup>9</sup> assumed that acoustic propagation at low frequencies follows a wave mode characterized by the conformal grid characteristic of streamline flow:<sup>10</sup> the wavefronts are defined by curved lines of constant-pressure in two dimensions (2Ds), curved surfaces in three dimensions (3Ds), that intersect the vocal-tract walls at right angles.<sup>8,9,11,12</sup> The flow in an acoustic wave is a local back-and-forth disturbance, not a bulk streaming flow, but it is assumed for the stated conditions that the acoustic particle velocities are aligned with streamlines under steady potential flow. Pressure measurements of cast models of the vocal tract<sup>13</sup> along with a finite-element model<sup>14</sup> (FEM) confirm those assumptions for frequencies up to 4 kHz for the shapes studied, although another FEM study showed a departure from such a plane-wave propagation mode at 1.5 kHz in the case of the large front cavity for */r/*.<sup>15</sup>

A technique called hyperbolic grid generation<sup>16–19</sup> is commonly used to subdivide a region bounded on one side in 2D or 3D into grid elements along curved paths for the purpose of solving a partial differential equation (PDE) on that region.<sup>20</sup> The PDE could be for heat transfer, fluid flow, or as is the case here, acoustic wave propagation; grid generation is not the solution of that PDE but rather a preparatory step. The generated grid also happens to be the solution to a hyperbolic PDE, one describing the propagation of a grid-generating wavefront. Differing from a finite-difference PDE solver that locates wavefronts by interpolating from solution values at fixed grid points, the hyperbolic method directly expresses the next location of the wavefront by placing grid points along vectors directed away from points on the current location of the wavefront. This so-called marching algorithm computes the grid in one pass, and is thus favored over other PDE-derived methods, such as those solving for the potential field and requiring iterative updates of the entire grid.

The hyperbolic algorithm from Henshaw<sup>18</sup> has been adapted to locate the combined lower lip, tongue, and anterior pharyngeal wall outline by marching in steps away from outer vocal tract outline in the 2D midsagittal plane. The total marching distance is determined by lip opening, the weighted sum of tongue-displacement basis functions, or a constant distance in the region superior to the vocal folds within the respective regions of the vocal tract. The grid lines between the outer and inner outlines are generated with each intermediate marching step. The automatic generation of the grid allows separating scaling of those basis functions from changes in the outline. The transverse grid lines generated by this process provide curved paths nearly perpendicular to the palate and tongue that more closely represent acoustic wavefronts. Substituting straight grid lines that intersect the palate and tongue at oblique angles is known to result in errors in computed formant frequencies.<sup>21</sup>

A dissipation term evens the distribution of transverse grid lines, and a marching-speed term retards the grid wave-

front when grid lines perpendicular to the wavefront are diverging and advances that wavefront when the perpendicular grid lines are converging. These two smoothing terms prevents the formation of grid shocks—crossings of the grid lines that result in sharp bends or cusps in the tongue shape for tongue displacements that exceed the radius of the local palate curvature. This allows the representation of a wide range of human-subject palate outlines rather than restricting the articulatory model to reference outlines with broad curvature.

The hyperbolic algorithm generates representations of acoustic streamline and wavefront paths as a byproduct of constructing the tongue shape. Keefe and Benade<sup>11</sup> calculated from streamline and wavefront paths a correction for the effective acoustic lengths and cross-sectional areas for a transmission-line model of strongly curved ducts. Resonance frequency shifts may be analytically determined for simple geometries, thus helping to explain the acoustic mechanism by which those shifts occur. This curvature correction is compared to the 3D wave mode solution from Sondhi's<sup>22</sup> analytical treatment of a curved duct. For this geometry as well as more complex ones, comparison is made to a finite-element analysis of the wave modes.

Hyperbolic grid generation coupled with a mathematical model of articulatory basis functions thus generates smooth tongue shapes representative of vowels in human speech. Given a model for the vocal-tract transverse cross section, grid lines connecting the palate to the tongue provide cross-distances for calculating the area function. This area function may in turn be corrected for vocal-tract curvature based on the streamline paths, and the intrinsic effect of curvature on formant frequencies may in this manner be quantified apart from the influence of other factors.

## II. METHODS

### A. Outer vocal-tract outline

Outer vocal-tract outlines were determined for two subjects, an adult male and an adult female, whose computed tomography (CT) x-ray images with superimposed tracings are shown in Fig. 1. These subjects received CT-scans for medical reasons not known to affect vocal-tract shape, but each subject was at rest and not speaking. As such, these data do not provide cross-distances or area functions, but distances and area functions representative of speech are supplied from published magnetic resonance imaging (MRI) data as described in Secs. II B–II F.

Each outline was traced from the midsagittal slice of a CT study using procedures described by Vorperian *et al.*<sup>23</sup> with changes noted below. These subjects received CT-scans for medical reasons not known to affect vocal-tract shape, and as such, the images are for a non-speech resting condition. A marker is drawn to locate a point on the anterior surface of the incisors at the level of the inferior surface of the upper lip. From that marker, a line is drawn to the lingual aspect of the central incisors. The reasons for interpolating through the incisors are (1) the teeth have gaps that permit acoustic wave propagation and (2) the volume displaced by

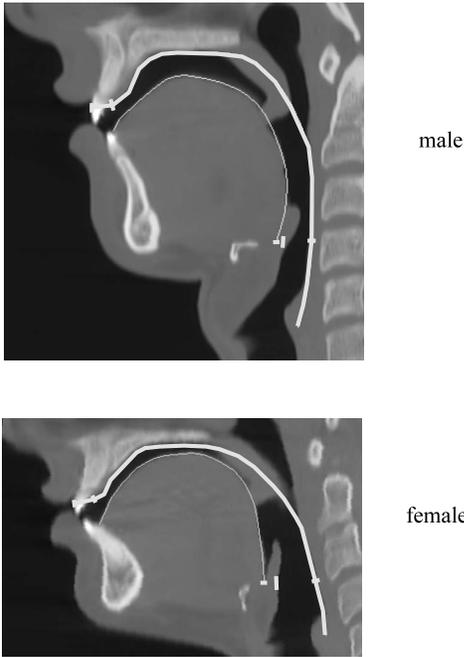


FIG. 1. Tracing of outer vocal-tract outline on CT-scan image.

the teeth represents a local perturbation in the area function affecting higher-order formants beyond those of interest.

A curvilinear path is continued by tracing along the length of the hard palate to the beginning of the soft palate. From there the curvilinear path crosses through the soft palate to the back of the pharyngeal wall down to the posterior aspect of the glottis. The path follows what is estimated to be the inferior surface of the soft palate when raised during vowel production. The demarcation between oral-pharyngeal and epilarynx regions was made with reference to the veliculae along with the superior surface of the hyoid. The epilarynx refers to the narrow conduit immediately superior to the vocal folds. The figure shows a mark for each subject denoting the separation of the anterior from the posterior boundaries of the epilarynx.

## B. Determining area function from articulatory parameters

Articulatory parameters specify the displacement of the inner vocal-tract outline from the outer vocal-tract outline. A saturation function for that displacement, given in Eq. (3) below, prevents collision of the tongue with the palate. Displacements at points along the outer outline are in turn converted into areas using a model for the vocal-tract cross section.

A weighted sum of front-raising, back-raising, and jaw-opening basis functions controls the outer outline-tongue distance in the oral-pharyngeal portion of the vocal tract by

$$d_b(x) = d_{\text{avg}}(1 - w_{\text{fr}}h_{\text{fr}}(x) - w_{\text{br}}h_{\text{br}}(x) + w_{\text{jo}}h_{\text{jo}}(x)), \quad (1)$$

where average opening  $d_{\text{avg}}=1.0$  cm,  $x$  is longitudinal distance along the central vocal-tract region normalized to  $0 \leq x \leq 1$ , the  $w$  coefficients vary with articulation, and setting tongue phase  $\theta_t(x)=(1.75x+0.25)\pi$  and jaw phase  $\theta_j(x)$

$= (1.3x)\pi$  define the front-raising (fr), back-raising (br), and jaw-opening (jo) basis functions:

$$h_{\text{fr}} = \sin \theta_t, \quad h_{\text{br}} = -\cos \theta_t, \quad h_{\text{jo}} = 0.3 + 0.7 \cos \theta_j. \quad (2)$$

Liljencrants<sup>7</sup> originally proposed the Fourier basis for the outline-tongue distance, but the first order Fourier terms produced shorter constriction regions than x-ray observations, requiring the addition of higher order Fourier terms. A stretching of the first order Fourier terms gives longer constriction regions. For example, basis function  $h_{\text{fr}}$  places the constriction of /i/ forward in the vocal-tract compared to the un-stretched Fourier basis function  $\sin(2\pi x)$  while keeping the shape of the back cavity the same. A linear weighting of the front and back raising basis functions controls the place and depth of the tongue constriction after  $A \cos x + B \sin x = \sqrt{A^2 + B^2} \cos(x + \arctan(B/A))$ .

The jaw-opening function has a longer wavelength than the front raising function and the combination of a positive jaw weight with a countervailing front raising weight produces the flared horn-like shape of /æ/. The jaw-opening function has a larger front opening relative to back narrowing; varying the jaw opening helps match the variation in the vocal-tract volume observed in MRI data, as will be discussed in Sec. II F.

A displacement saturation function imposes a minimum 0.15 cm vocal-tract opening according to

$$d_{\text{vt}} = 0.0505e^{4d_b} + 0.15, \quad d_b < 0.4,$$

$$d_{\text{vt}} = d_b, \quad d_b \geq 0.4, \quad (3)$$

where basis-function distance  $d_b$  and corrected vocal-tract distance  $d_{\text{vt}}$  are in centimeters. As the vocal tract is constricted, this saturation effect starts at a cross-distance of 0.4 cm, is continuous in both the function and its derivative between the linear and exponential portions of the curve, and has a limiting minimum distance of 0.15 cm for extreme negative undershoot of the value of  $d_b$ .

The longitudinal length of the lip region has been fixed at 1 cm for both the male and female subjects, and the lip-opening parameter controls the midsagittal cross-distance. The length and cross-distance of the epilarynx region are fixed by measurements of the CT image. The lip-vocal tract and vocal tract-epilarynx boundaries are given transition regions of 1 cm length.

The proposed transverse cross section for the central vocal tract is an inverted parabola with the apex at the palate, which is a simplified, flat-bottomed version of a profile attributed by Perrier *et al.*<sup>24</sup> to Maeda.<sup>25</sup> A parabola with palate-tongue distance 1 cm and width 3 cm is representative of CT transverse cross section images of the male subject, giving area from palate-tongue distance in  $\text{cm}^2$  as

$$A = \alpha d_{\text{vt}}^\beta, \quad (4)$$

where  $\alpha=2$  and  $\beta=1.5$ . This  $\beta$  is for a parabola whereas the  $\alpha$  value is within reported ranges.<sup>24,26</sup> The model lip cross section is a pair of parabolas with bases meeting in the middle of the lip opening, scaled to give the same formula for area in terms of total lip opening to reduce interpolation artifact. The epilarynx cross section is an ellipse where the

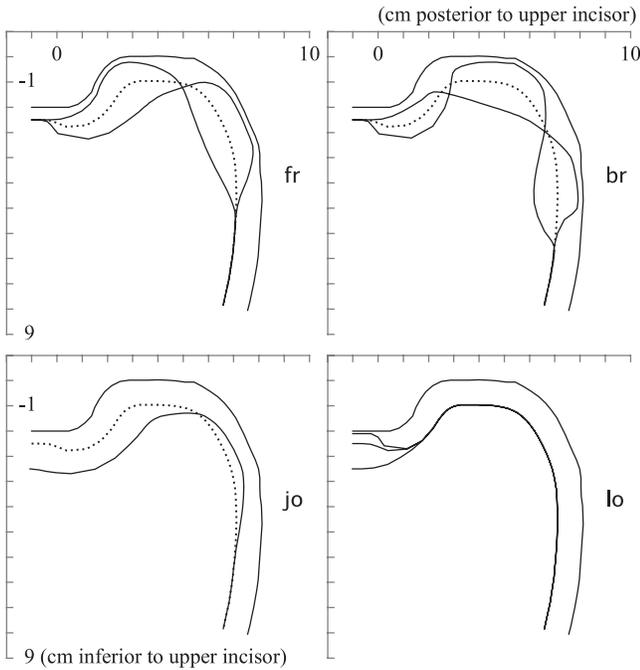


FIG. 2. Variation of the front-raising, back-raising, jaw-opening, and lip-opening basis-function weights for the male vocal-tract outline viewed in the midsagittal plane.

major axis is the anterior-posterior distance of the opening in that region and the minor axis is half that distance, giving area

$$A_{\text{epi}} = \frac{\pi}{8} d^2. \quad (5)$$

Cross-distance in the transition regions is determined by

$$d_{\text{trans}} = d_{\text{vt}} \cos \frac{\pi x_t}{2} + d \left( 1 - \cos \frac{\pi x_t}{2} \right), \quad (6)$$

where  $x_t$  ranges from 1 to 0 in the lip-vocal tract transition and from 0 to 1 in the vocal tract-epilarynx transitions so as to bias the distances toward the values at the ends of the vocal tract. Area is calculated in the transition region by linear blending of area formulas for neighboring regions.

The tongue contour changes after adjusting individual basis-function weights are shown in Fig. 2 for the male vocal-tract outline. The tongue displacement is along a grid generated by the hyperbolic method presented in Sec. II C.

### C. Hyperbolic grid generation

Grid  $\mathbf{x}(s, t) = (x_x(s, t), x_y(s, t))$  denotes a collection of  $x$ - $y$  points in the midsagittal plane. Integer index  $0 \leq s \leq N$  selects placement along the outer vocal-tract outline ( $s=0$  at the lips,  $s=N=50$  at the glottis for a 50-section vocal tract), and  $0 \leq t \leq 1$  denotes normalized distance between the outer and inner vocal-tract outlines. Curved paths connecting  $x$ - $y$  points of constant  $t$  in this grid represent streamlines; paths of constant  $s$  represent acoustic wavefronts. The grid function  $\mathbf{x}(s, t)$  thus maps a rectilinear grid in  $t$ - $s$  space into a curved grid in  $x$ - $y$  space; variable  $t$  also acts like time in the progression of the grid-generating wavefront from outer to inner vocal-tract outlines, propagating in a direction trans-

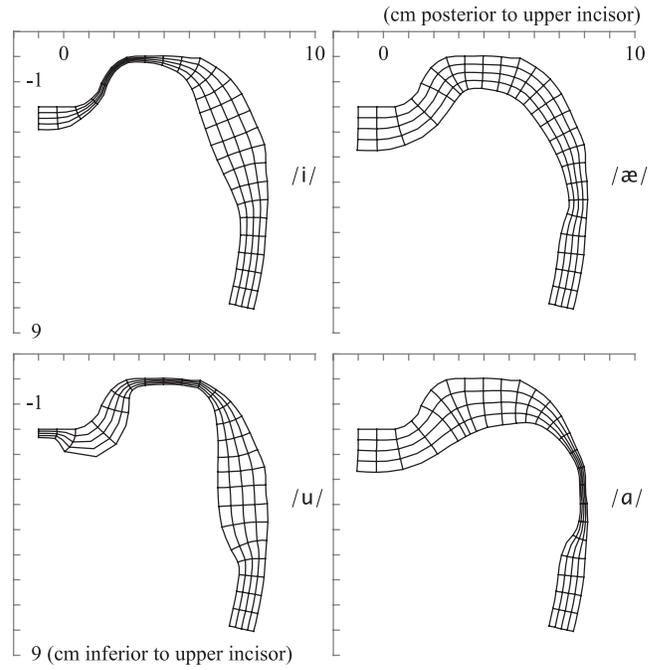


FIG. 3. Midsagittal vocal-tract shapes of vowels for the male vocal-tract outline generated by matching the articulatory model area function to MRI area functions from Story *et al.* (Ref. 35).

verse to the acoustic wavefront. Figure 3 shows generated grids of cardinal vowels: only every other grid line is plotted for clarity.

The partial derivative of grid  $\mathbf{x}(s, t)$  with respect to parameter  $t$  is given by

$$\begin{aligned} \mathbf{x}_t(s, t) = & S(s, t) \mathbf{n}(s, t) + \varepsilon_d \mathbf{x}_{ss}(s, t) \\ & - (\mathbf{x}_{ss}(s, t) \cdot \mathbf{n}(s, t)) \mathbf{n}(s, t), \end{aligned} \quad (7)$$

where  $S(s, t)$  is the local speed of the grid-generating wavefront, wavefront unit normal vectors  $\mathbf{n}(s, t)$  are computed from tangent slopes derived from a fit of piecewise circular arcs to points on that wavefront,<sup>27-29</sup> second central difference  $\mathbf{x}_{ss}(s, t) = \mathbf{x}(s-1, t) - 2\mathbf{x}(s, t) + \mathbf{x}(s+1, t)$ , and term  $\mathbf{x}_{ss}(s, t) - (\mathbf{x}_{ss}(s, t) \cdot \mathbf{n}(s, t)) \mathbf{n}(s, t)$  denotes that part of  $x$ - $y$  vector  $\mathbf{x}_{ss}(s, t)$  normal to  $\mathbf{n}(s, t)$ . A dissipation coefficient is scaled in relation to the time step according to  $\varepsilon_d = 0.4dt$ , smoothing the distribution of points along the wavefront path to suppress crowding of the grid lines intersecting the wavefront by allowing those lines to deviate from being precisely normal to that wavefront.

Equation (7) is solved using the second-order Runge-Kutta numerical integration of Petersson with more steps<sup>19</sup> in place of the implicit integration of Henshaw<sup>18</sup> having more calculations per step. The integration divides the displacement between palate and tongue into steps  $dt$ , where  $dt=1$  would apply the entire displacement from palate to tongue in one step. Employing a multi-step marching algorithm in place of a single-step procedure avoids crossing of grid lines, a problem when the palate-to-tongue distance exceeds the local radius of the curved palate. In addition to the aforementioned dissipation term, adjustments to the marching speed  $S$  can suppress crossing of the grid lines while providing control over the palate-tongue distance. Whereas the speed ad-

justments used to prevent grid-line intersection also prevent an exact match to a target palate-tongue distance, the following formula keeps the effect of distance mismatch small relative to other effects being investigated:

$$S(s,t) = d(s,t) \frac{\overline{\Delta a}}{\Delta a} (1 - \varepsilon_c(t) d(s,t) \kappa(s,t)). \quad (8)$$

The terms of this equation are as follows, starting with  $d(s,t)$  computed for each marching step from an initial palate-tongue distance  $d(s,0)$ , supplied by the weighted vocal-tract basis functions, using the update

$$d(s,t_k) = d(s,t_{k-1}) \left( 1 + \frac{\alpha}{8} \right) - \alpha |\mathbf{x}(s,t_k) - \mathbf{x}(s,t_{k-1})|, \quad (9)$$

which compensates for changes in the palate-tongue distance brought about by the smoothing terms of  $S(s,t)$ , where  $t_k = k/8$  and  $\alpha = (1 - \sqrt{t_{k-1}})/(1 - t_k)$  for  $0 < k < 8$ . Choice of  $\alpha$  controls the trade-off between a smooth tongue surface and a precise match to the basis function-determined palate-tongue distance.

Defining backward and forward differences  $\mathbf{x}_{s-} = \mathbf{x}(s,t) - \mathbf{x}(s-1,t)$  and  $\mathbf{x}_{s+} = \mathbf{x}(s,t) - \mathbf{x}(s+1,t)$ , speed term  $\Delta a / \overline{\Delta a}$  adjusts marching distances to smooth out local values of wavefront arc length  $\overline{\Delta a} = 0.5(|\mathbf{x}_{s-}| + |\mathbf{x}_{s+}|)$  that deviate from averaged arc length  $\Delta a$  where the averaging is performed by eight applications of the difference equation,

$$\Delta a_1(s) = 0.8375 \Delta a(s) + 0.08125 (\Delta a(s-1) + \Delta a(s+1)). \quad (10)$$

Finally, speed term  $1 - \varepsilon_c(t) d(s,t) \kappa(s,t)$  adjusts the marching steps to smooth local wavefront curvature

$$\kappa(s,t) = \frac{\mathbf{n} \cdot \mathbf{x}_{s-}}{|\mathbf{x}_{s-}|^2} + \frac{\mathbf{n} \cdot \mathbf{x}_{s+}}{|\mathbf{x}_{s+}|^2}. \quad (11)$$

Curvature is the reciprocal of the local radius of the wavefront; the product of curvature with the palate-tongue distance  $d(s,t)$  is thus dimensionless as required for the formula. The coefficient of the curvature adjustment

$$\varepsilon_c(t) = 0.183 \sqrt{t(1-t)} \quad (12)$$

has a maximum value midway between palate and tongue and minimum values at the ends to give maximum smoothing of the streamline paths midway between these two boundaries. This amount of curvature adjustment prevented the intersection of palate-tongue grid lines for the vocal-tract outlines and range of tongue displacements under consideration.

Too large a step size  $dt$  relative to the wavefront arc length  $\Delta a$  results in the growth of local disturbances. A bound on the step size that suppresses such disturbances is expressed as

$$dt = \min_s \left\{ \left( \frac{1}{2\varepsilon_c} \right) \left( \frac{|\mathbf{x}_{s-}|^2 + |\mathbf{x}_{s+}|^2}{2d(s,t)^2} \right) (\Delta a / \overline{\Delta a}) \right\}. \quad (13)$$

## D. Curvature correction of acoustic-tube lengths and areas

The vocal tract in vowel production is modeled by a non-uniform acoustic tube, where the cross sectional area is a function of distance along a curved path from lips to glottis. A curvature correction for a transmission-line analog after the method of Keefe and Benade<sup>11</sup> follows and is compared with the 3D wave mode solution of Sondhi and a 3D FEM solution.

A curvature correction to the lengths and areas of elements of the non-uniform acoustic tube may be understood by considering that a volume element of the duct that has length  $l$  in the direction of fluid motion, cross-sectional area  $A$ , and nearly constant-pressure and flow within that volume has acoustic inductance:

$$L = \frac{\rho l}{A}, \quad (14)$$

where  $\rho$  is the density of air and has capacitance

$$C = \frac{Al}{\rho c^2}, \quad (15)$$

and where  $c$  is the speed of sound. In turn, an effective acoustic length  $l$  and area  $A$  may be determined from specified values of  $L$  and  $C$  according to

$$l = c \sqrt{LC}, \quad A = \rho c \sqrt{\frac{C}{L}}. \quad (16)$$

A short section of a duct between a pair of constant-pressure surfaces may be divided into longitudinal sections representing stream tubes. Figure 3 of Keefe and Benade<sup>11</sup> illustrates a duct section and one of these stream tubes. The bundle of stream tubes sees a constant-pressure at each face, and the difference in pressure between the two faces acts on the fluid inertia to change the rate of flow in each stream tube. From circuit theory, the capacitances of these parallel circuit branches add, and the reciprocal inductance is the sum of the reciprocal inductances. The branch capacitance and inductance can be determined from the average area  $A$  and average length  $l$  of each stream tube, the capacitances and inductances of those branches can be combined, and the effective acoustic area  $A$  and length  $l$  of the duct section can be solved from the combined  $L$  and  $C$  values.

For a curved duct of constant cross section and bend radius, the constant-pressure surfaces are assumed to be radial slices, and the streamlines to follow circular paths between the inner and outer bend radius of the duct. The spacing of the streamlines between the inner and outer bend radius is not required; the only required assumption is that in a duct with a circular bend, geometric symmetry has the streamlines following circular paths. Wave mode solutions<sup>22,30</sup> show pressure variation in the radial direction, even for the plane-wave mode. Keefe and Benade<sup>11</sup> considered radial pressure variation of the form expected at low frequencies and did not find a significant change from constant-pressure. Keefe and Benade,<sup>11</sup> however, found that resonant frequencies of a bent tube measured in the laboratory changed in the same direction but in a lower amount

TABLE I. Formant frequencies in Hz for a curved vocal tract of uniform cross section, rigid walls, and zero wall and radiation loss. Curvature corrected: adjusts area and length by combining stream tubes; modal: acoustic wave mode solution of Sondhi (Ref. 22), FEM: 3D finite-element analysis.

	Rectangular cross section		Parabolic cross section		
	Curvature corrected	Modal	FEM	Curvature corrected	FEM
F1	502.9	502.7	503.0	521.8	522.0
F2	1508.7	1504.4	1505.4	1565.3	1563.3
F3	2514.6	2494.9	2497.6	2608.9	2596.9

than the prediction from integration over stream tubes. Nederveen<sup>12</sup> attributed this difference to small errors in measuring tube diameters in brasswind musical instruments; controlling for tube diameter gave close agreement between acoustic measurements and stream tube theory.

The curvature correction to the transmission-line analog may be compared to the 3D wave mode solution of the acoustic wave equation  $\nabla^2 p = (1/c^2)(\partial^2 p / \partial t^2)$ , where  $p$  is time and spatially varying pressure and  $c$  is the speed of sound. Separation of the time variable results in the Helmholtz equation  $\nabla^2 P = -(\omega^2/c^2)P$ , where  $P$  is the spatially varying coefficient of the sinusoidal time dependence of pressure. Sondhi<sup>22</sup> solved this equation analytically for a curved duct with hard walls, a radiation model of a simple acoustic short circuit, a constant bend radius, and a rectangular cross section.

The finite-element approximation to the wave mode is the solution to the matrix equation  $KP = \lambda MP$ .<sup>31</sup> Acoustic pressure in this approximation is assumed to vary linearly within tetrahedral volume elements between node values at the corners of those elements. For a lossless duct, this matrix equation is a generalized eigenvalue problem where matrices  $K$  and  $M$  are symmetric and positive definite, giving real-valued solutions for scalar  $\lambda = \omega^2/c^2$  relating to mode frequency and vector  $P$  of node pressures relating to mode shape. A JAVA program calculates coefficients of the matrices  $K$  and  $M$  from the coordinates of the corners of the finite elements using the formulas from Huebner,<sup>31</sup> and the matrix equation is solved using the DSBGV double-precision symmetric-banded generalized eigenvalue, function from the LAPACK software library.

Table I presents results for a curved duct with a 17 cm midline length,  $z$ -axis width 2.5 cm, forming a 90° arc having cross-distance of 4 cm and bend radii  $r_1 = 8.82$  cm,  $r_2 = 12.82$  cm. Using Sondhi's value for sound speed of 34 000 cm/s,<sup>22</sup> an equivalent straight duct has formants F1, F2, and F3 at 500, 1500, and 2500 Hz. Table I reports results for the preceding curvature correction, Sondhi's wave mode analysis, and a 3D FEM. The duct is assumed throughout to have rigid walls along with zero wall and radiation loss.

For constant-pressure across radial slices, integration of a duct of rectangular cross section and inner radius  $r_1$  and outer radius  $r_2$  determines an effective inductance,

$$L_{\text{eff}} = \frac{r_2 - r_1}{\ln(r_2/r_1)(r_2 + r_1)/2} L, \quad (17)$$

where  $L$  is the uncorrected inductance calculated from the midline distance; the capacitance  $C$  remains unchanged, and effective tube length and area may be calculated using Eq. (16). Table I also includes results for a parabolic cross section of height 4 cm and width 3.75 cm, having the same area as the rectangular cross section. The integrations for the curvature correction for this parabolic cross section were carried out with the MAPLE symbolic solver software package.

This finite-element analysis divides the duct into 50 hexahedral (six-sided box-like) sections along the longitudinal axis, eight sections in the radial direction, and one section along the  $z$ -axis normal to the plane of the bend. Increasing the number of hexahedra in the  $z$ -direction to 4 resulted in a formant shift of less than 0.3 Hz. Each hexahedral section is subdivided into six tetrahedral elements in such a way that neighboring pairs of tetrahedra share a face without overlap into a third tetrahedron: this condition ensures that pressure is piecewise linear throughout the entire duct volume without any step changes. Increase in the number of sections along the longitudinal and radial directions lowered the formant frequencies for the FEM in the direction of Sondhi's analytical modal results.

The curvature correction to the transmission-line model predicts a linear scaling of formant shift with increasing formant frequency. The modal and FEM values agree with the curvature correction at F1, but those values vary downward from linear scaling for higher formants. This decline may be related to the influence of the cross modes on the lower plane-wave mode resonances.<sup>32</sup> Sondhi's tables show the amount of decline to be reduced for a smaller bend radius difference, a change that raises the frequency of the first cross mode. The decline results in a net downward shift for F3 for the rectangular cross section not predicted by the curvature correction, but this effect is masked by a much larger upward shift in formants intrinsic to the curvature correction for a parabolic cross section

Applying the curvature correction to the area function of more complex vocal-tract shapes representative of vowel production requires numerical methods. A hyperbolic grid, combined with the simplified version of Maeda's parabolic model for the transverse cross section, divides the vocal tract into stream tube sections. The span between palate and tongue is once more divided into eight stream tubes, and the longitudinal axis from lips to glottis is divided into 50 sections. The cross section of each end of a stream tube section is approximated with a trapezoid of the same area. The distance between centroids of the trapezoid at each face determines the section length. The volume of each stream tube is determined by adding the volumes of a spatial tiling of tetrahedra as done for the FEM, and area of a stream tube section is calculated as volume divided by length.

The vocal tract is represented as a series of uniform-tube sections, and section values for  $1/L$  and  $C$ , and hence the effective length and effective area, are determined by summing over  $1/L$  and  $C$  values, respectively, for the discrete stream tubes marked by the grid. Numerical determination of

formant frequencies from a vocal tract represented by a series connection of uniform tubes of specified length and cross-sectional area is addressed in Sec. II E.

### E. Determination of formant frequencies from area function

The method for calculating the acoustic frequency response of a series connection of acoustic tubes is based on the frequency-domain method of Atal *et al.*,<sup>33</sup> which takes into account the effects of viscous and heat conduction losses within the vocal-tract along with the yielding vocal-tract sidewall. The frequency-dependent lip radiation load is computed using a formula for a circular tube terminating in an infinite-plane baffle.<sup>34</sup>

Formant frequencies may be estimated by evaluating vocal-tract transfer function magnitude  $|H|$  at different values of  $\omega$  to determine peaks. This commonly used method encounters the problem of merged peaks, occurring when the spacing of a pair of formants is small relative to their bandwidths. Atal *et al.*<sup>33</sup> addressed this by evaluating  $|H|$  at complex values away from the imaginary axis  $s=j\omega$  and by finding zeroes of  $1/H$  by a Newton-Raphson search procedure. An alternative formant-finding method follows from  $1/H$  being real-valued in the limiting case of a lossless vocal tract. For a lossy vocal tract, the real part of  $1/H$  is evaluated for zero crossings, which are empirically determined to be close to the frequencies of peaks of  $H$  but provide better resolution of closely spaced formants. Once a zero crossing is found by evaluating  $1/H$  at discrete steps  $\Delta\omega=2\pi 40$  Hz, interval bisection and linear interpolation refine the location of that zero.

The frequency-domain method was evaluated using MRI-derived area functions from Story *et al.*<sup>35</sup> That study computed formant frequencies by a different method, one employing a wave-analog synthesizer incorporating a side branch for the piriform sinuses. In place of this side branch, the frequency-domain method was modified by adding volume to the acoustic tube section immediately superior to the epilarynx. Dang and Honda<sup>36</sup> reported total piriform volume for an adult subject in excess of 2 cc. A frequency-domain calculation using a piriform volume correction of 1.25 cc, wall mass coefficient of 1 gm/cm<sup>3</sup>, and sound speed of 35 000 cm/s gave the closest match ( $\Delta F_2$  of /u/ <60 Hz,  $\Delta F$  all others <20 Hz) to formant F1-F2 values reported by Story *et al.*<sup>35</sup>

### F. Matching model to MRI area functions

Weights for lip opening along with three tongue basis functions were manually adjusted to obtain a best compromise between matching model-generated to the MRI-measured area functions from Story *et al.*<sup>35</sup> and matching the first two formant frequencies observed as peaks of the frequency response. This procedure determined basis function weights representative of the vowels /a/, /æ/, /i/, and /u/, obtaining a single set of articulatory parameter values to substitute into different vocal-tract profiles. These matches were obtained using a computer-generated display of the midsagittal plane vocal-tract profile, the area function derived from

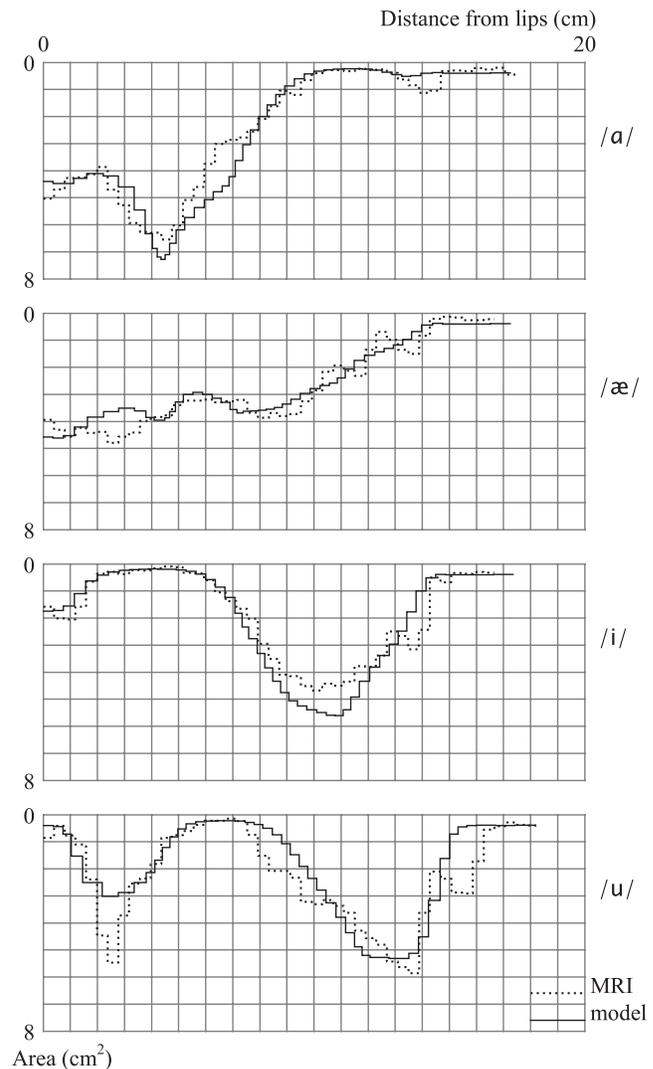


FIG. 4. Results of matching articulatory model area functions to MRI area functions from Story *et al.* (Ref. 35).

that profile, and the acoustic frequency response. Acoustic-tube lengths were determined from the longitudinal grid midline, without the curvature correction, to best represent the way the target MRI area functions had been measured.

The resulting midsagittal outlines are shown in Fig. 3; Fig. 4 shows the match between area functions. These results proved to be sensitive to the length of the epilarynx tube. The male subject epilarynx tube was shortened by 0.35 cm from the CT-scan-derived area functions shown in Fig. 4 prior to calculating formants. This adjustment resulted in formant frequencies between model-generated and MRI area functions that were within  $\Delta F_1$  of /a/ <25 Hz,  $\Delta F$  all others <10 Hz.

The effects of varying the epilarynx length and adding piriform sinus volume are seen in Fig. 5. The reference condition is for the epilarynx length measured from the CT-scan without piriform sinuses. Shortening the epilarynx in the amount of 0.35 cm, the condition in the match to the MRI data, raised  $F_2$  for /i/. Keeping the measured epilarynx length and instead adding a 1.25 cc piriform volume lowered  $F_1$  and  $F_2$  for /a/ and /æ/.

MRI (Refs. 37 and 38) and CT (Ref. 24) studies of the

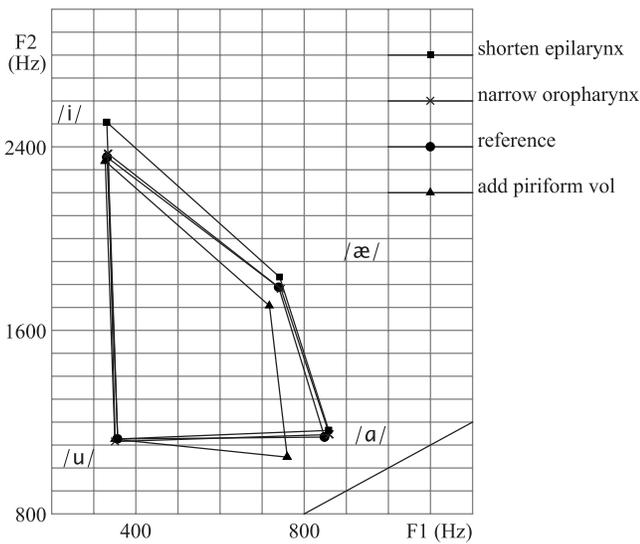


FIG. 5. Formant sensitivity to changes in the lower vocal tract: the reference condition is for the male outline with no piriform sinuses, the shortened epilarynx condition is for a length reduction of 0.35 cm, and the added piriform volume condition adds 1.25 cc to the piriform sinuses with the epilarynx length restored to the reference.

coefficients of the  $\alpha d^\beta$  model for computing area from cross-distance  $d$  show variations in these coefficients with place in the vocal tract. Two of these studies<sup>24,38</sup> show a consistent  $z$ -axis narrowing of the vocal tract in the oropharynx, related to a reduction in  $\alpha$  or in  $\beta$ . The “narrow oropharynx” condition in Fig. 5 relates to a raised cosine function spanning 4 cm in length, located midway in the central vocal-tract region, scaled to give a peak reduction in  $\alpha$  of 20%. The deviation from the reference condition resulting from this perturbation is small relative to the effects of variation in the epilarynx length and piriform volume.

### III. RESULTS

#### A. Comparison of methods for determining area function

The effect of the curved vocal tract on formant frequencies is evaluated by comparing four methods for computing the vocal tract area function from the basis function model. This comparison is made on the male-subject outline with the CT-scan measured epilarynx length, using the basis function weights for the /a/, /æ/, /i/, and /u/ obtained by matching MRI data; the results are shown in Figs. 6–8. Numerical values of formants from the four methods are presented in Table II.

The “palate lengths” method refers to an unrolled vocal tract where the weighted combination of basis functions directly specifies palate-tongue cross-distances used to compute acoustic-tube area, where acoustic-tube lengths are determined with reference to distances along the outer vocal-tract outline. The “grid smoothed” method also uses palate-referenced acoustic-tube lengths, but the cross-distances are taken from the generated grids shown in Fig. 3. The “midline lengths” method uses the smoothed grid cross-distances for area, but distances along the grid midline extending from lips to vocal folds give acoustic-tube lengths. Finally, the “curva-

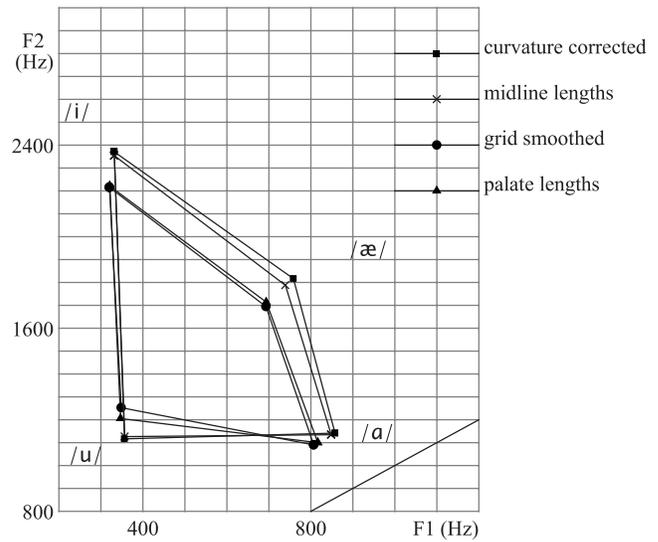


FIG. 6. Formant frequencies determined for the male vocal-tract outline using four methods of determining acoustic-tube lengths and areas.

ture corrected” method makes adjustments to acoustic-tube areas and lengths by combining stream tubes.

The formant shifts brought about by grid smoothing are small: this is a result of the choice of grid smoothing coefficients along with the adjustments made to grid marching speed to compensate for changes in cross-distance brought about by grid smoothing. Some shift in the cross-distances needs to be tolerated to prevent crossing of grid lines and the resulting sharp cusps in the tongue, but the grid-generation algorithm has been tuned so that the computed formant shifts that are small relative to the other effects under consideration.

Switching the acoustic-tube lengths from distances along the palate to the grid-generated midline brought about the greatest change in formants. Formant frequencies increased as a result of shorter midline distances and hence a shorter overall vocal-tract length relative to the palate distances. The curvature correction resulted in a small addi-

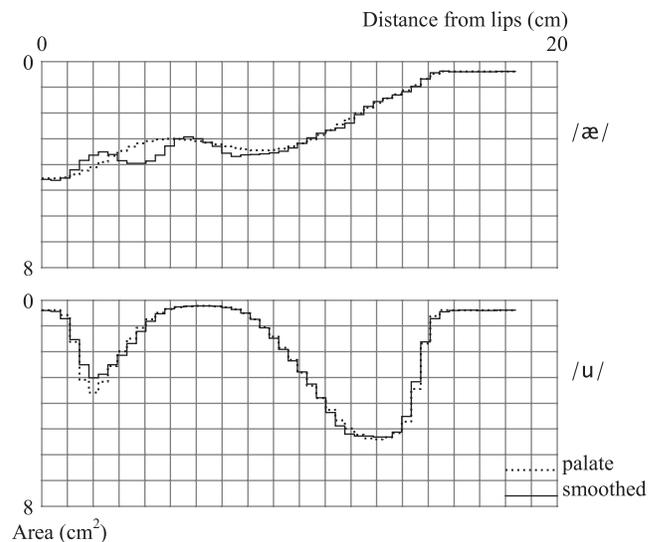


FIG. 7. Change to area function resulting from grid smoothing.

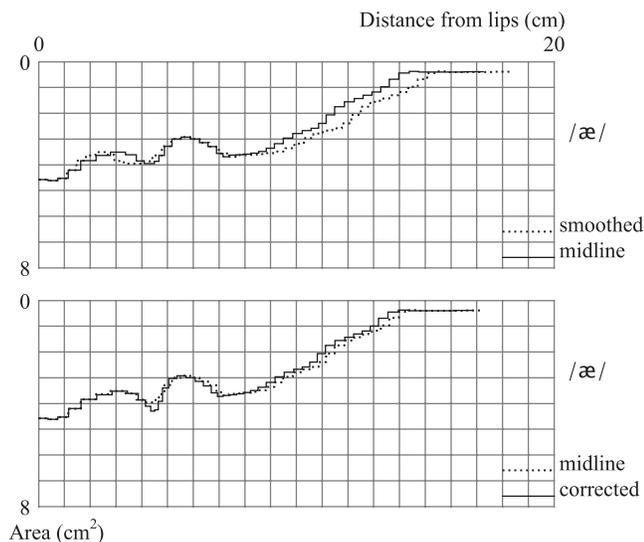


FIG. 8. Change to area function resulting from midline-distance correction, curvature correction.

tional increase in frequency relative to that midline-distance effect, an amount consistent with the analysis of simple ducts in an earlier section.

Consequences of the interaction of grid smoothing with vocal-tract curvature on the area function are seen in Fig. 7 for /æ/ and /u/. The vowel /æ/ has the simplest geometry of a horn-like flared tube while /u/ has the most complex geometry, with front and back cavities separated by a central constriction. Grid smoothing keeps the same tube lengths while producing area changes, which are greatest in the front part of the vocal tract having the reverse curve between the convex alveolar region and the concave palate vault. The area changes were larger for /æ/, where the grid algorithm had to march a greater distance from the palate, but the formant shifts were greater for /u/, where F2 is sensitive to changes in the front cavity.

The effect of changing from outer vocal-tract outline (palate) to midline-referenced acoustic-tube lengths for the grid-smoothed area function followed by changing to the

curvature correction for /æ/ are seen in Fig. 8. The switch from palate to midline distances keeps the same acoustic-tube areas but changes tube lengths in proportion to the midline stretching in convex regions and shrinking in concave regions. The curvature correction results in changes to vocal-tract area, but the changes in length are small compared to the palate-to-midline shift, confirming the effect seen in simple ducts.

Table III compares the curvature correction of the transmission-line model with FEM analysis of the wave modes, where the hyperbolic grid gives the hexahedral sections. Values are reported as the change in formant frequency from the reference condition where acoustic-tube lengths in the transmission-line model come from midline lengths. This comparison is conducted for a lossless vocal tract; as a consequence, the values for the reference midline condition differ from those reported in Table II.

The FE1 condition uses a single layer of hexahedral sections in the  $z$ -direction normal to the midsagittal plane; the FE4 condition uses four layers. The curvature correction is in agreement with the FE1 condition to within 2 Hz for F1, 6 Hz for F2, and 9 Hz for F3. The FE4 condition results in much larger change from FE1 than the 0.3 Hz noted for the curved duct of uniform cross section. These shifts may relate to the convergence and divergence of the vocal tract in the  $z$ -direction resulting from variation in cross-sectional area along with the change in transverse profile between the elliptic epilarynx tube and the parabolic oral-pharynx region. Such variation appears to require consideration of the 3D compound curvature of the acoustic wavefront; the proposed curvature correction along with the FE1 condition only account for the curvature in the midsagittal plane.

The FE4 condition involves solution of an eigenvalue problem for 2295 node variables, a computation taking 64 s on a 1.2 GHz Pentium III processor. The transmission-line calculation of frequency response takes 20 ms, about a factor 3000 increase in speed, which becomes important when large numbers of vocal-tract shapes need to be computed.

TABLE II. Formant frequencies in Hz for the male vocal-tract outline using the four methods of determining acoustic-tube lengths and areas reported in Fig. 6.

		Palate lengths	Grid smoothed	Midline lengths	Curvature corrected
/a/	F1	817.3	806.8	849.1	857.8
	F2	1099.4	1090.2	1134.1	1141.1
	F3	2627.9	2624.6	2956.4	3029.0
/æ/	F1	694.2	693.6	740.1	758.6
	F2	1713.5	1694.0	1786.7	1815.8
	F3	2312.6	2298.7	2414.5	2461.1
/i/	F1	322.0	320.8	332.1	332.6
	F2	2220.3	2215.0	2353.1	2370.6
	F3	2879.8	2873.3	3109.5	3141.6
/u/	F1	347.3	349.1	357.3	356.5
	F2	1205.1	1251.4	1125.7	1115.7
	F3	2542.3	2572.6	2701.5	2720.4

TABLE III. Formant frequencies in Hz for the male vocal tract outline computed under lossless condition for comparison with the finite-element method. Comparisons are given as frequency differences with the midline lengths condition. Curvature corrected: adjusts area and length by combining stream tubes; FE1: finite-element method with one hexahedral (box) section in the  $z$ -direction; FE4: four hexahedral sections in the  $z$ -direction.

		Frequency		Frequency difference from midline lengths		
		Midline lengths	Curvature corrected	FE1	FE4	
/a/	F1	829.7	6.9	6.8	3.6	
	F2	1154.7	12.7	12.9	9.3	
	F3	3082.5	69.6	72.6	52.5	
/æ/	F1	729.9	23.4	22.9	21.5	
	F2	1832.9	29.7	30.1	16.1	
	F3	2521.1	51.8	50.8	35.2	
/i/	F1	223.9	1.2	2.9	2.3	
	F2	2337.1	17.3	18.3	-18.6	
	F3	3155.6	38.8	47.5	27.5	
/u/	F1	260.4	-0.3	1.4	0.3	
	F2	1209.9	-12.8	-6.8	-12.6	
	F3	2688.3	19.6	23.6	-17.9	

## B. Effects of oral-pharyngeal length ratio and bend radius

Results for changing the outer vocal-tract outline while keeping the same basis functions and weights to generate the tongue shape are presented in Figs. 9 and 10. Four vocal-tract outlines are compared. To control for vocal-tract length effects, these outlines are all normalized to have the male subject outer vocal-tract outline length, also matching lip length along with length and cross-distance of the epilarynx

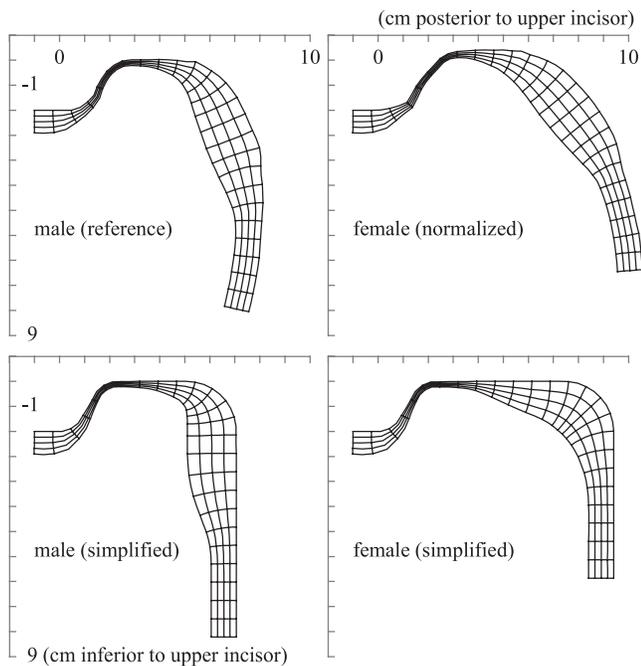


FIG. 9. Alternative vocal-tract shapes for vowel /i/ for different outer vocal-tract outlines. The natural male subject outline is the reference to which the other outlines are length normalized. The simplified shapes are stylized representations of the male and female shapes.

tube. The normalized female profile in Fig. 9 is the female profile from Fig. 1 after applying this process.

The effects of reduced bend radius are evaluated using stylized or simplified vocal tracts that have a convex alveolar surface and a concave palate vault of radius 1 cm, a 60° incline joining the alveolar and palate regions, a 2-cm rise to the top of the palate vault, and a 2-cm radius of the curve joining the oral and pharyngeal sections. The oral length is from anterior surface of the incisors to the posterior pharyngeal wall; the pharyngeal length is from the superior extent of the palate to the superior end of the epilarynx. The simplified female outline has an oral over pharyngeal length

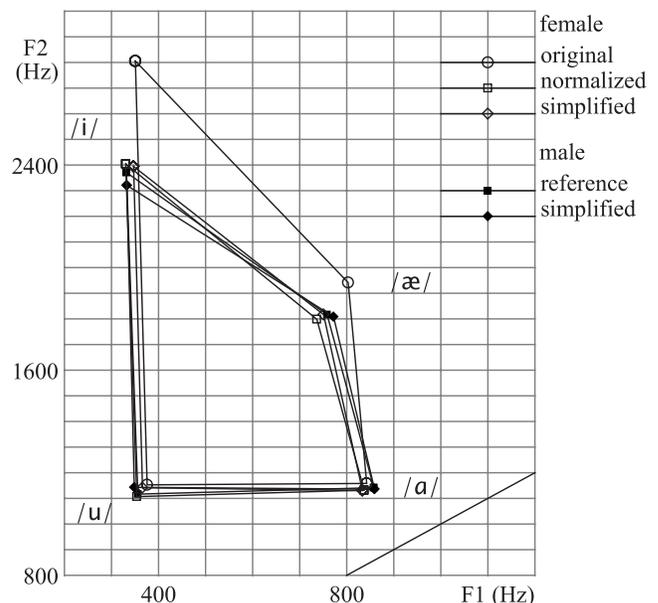


FIG. 10. Comparing model-derived formant frequencies from female outer vocal-tract outline with four outlines (female and male, natural and simplified) that are length normalized to the male outline.

ratio of 2.0 whereas the simplified male has a ratio of 1.0. Taking into account differences in degree and distribution of outline curvature, these simplified vocal tracts have similar proportions to the natural male and female outlines.

Figure 10 offers a comparison between formants computed using the curvature-correction method for the female subject before length normalization and the four length-normalized outlines. The four length-normalized outlines give tightly clustered formants compared to the male-female difference, confirming that length normalization can remove most of the difference in formant frequencies between the male and female subjects, and that differences in the placement and radii of vocal-tract bends provide a second-order effect.

#### IV. CONCLUSIONS

The goal of this work was to understand the mechanisms by which the vocal-tract bend causes shifts in formant frequencies. Methods based on manual placement of a rectilinear-radial grid have limitations relating to the fixed scaling of tongue displacement in relation to the grid, the requirement for a reference subject without a sharply curved palate, and the representation of acoustic wavefronts with straight lines oblique to the palate and tongue. The hyperbolic grid-generation algorithm addresses these limitations.

The hyperbolic method automatically generates a grid for positioning the tongue as a displacement from the palate that is adaptive to the vocal-tract shape, decoupling the scaling of the displacement profile from the location of fixed grid lines. The grid smoothing that takes place also allows tongue displacements that are large compared to the local radius of the palate, but that smoothing also contributes shifts to the area function that are influenced by vocal-tract curvature. Those shifts can be made small relative to other effects of the curved vocal tract by proper choice of smoothing coefficients. The resulting grid provides midline and cross-distances along curved paths for determining the lengths and areas of acoustic tubes for calculating formant frequencies; the grid also subdivides the space between palate and tongue in such a way as to allow a correction accounting for vocal tract curvature.

The dominant effect of the curved vocal tract on formant frequencies is the lengthening of the vocal-tract midline in the vicinity of convex curvature such as the alveolar region and the shortening of the midline for concave curvature such as the palate vault and the oral-pharyngeal bend. The curved vocal tract also produces a departure from linear formant scaling; vowels are affected differently because the midline length change is smaller for curvature in the neighborhood of a constriction and large when the separation between tongue surface and the outer vocal tract outline is large. The curved vocal tract results in pronounced outward expansion of the F1-F2 formant envelope for /æ/, where the primary effect of curvature is an overall shortening of the vocal tract. This pattern of envelope expansion stays consistent across changes in the vocal-tract cross sections.

The curvature correction for the transmission-line model is small relative to the changes attributable to the effect of

curvature on midline distances. Comparison with FEM along with consideration of acoustic sensitivity suggests that a more detailed treatment of 3D wave propagation effects is only significant if the vocal tract dimensions are known to high accuracy.

After normalizing for vocal-tract length, the amount the midline-distance effect varied with change in oral-pharyngeal length ratio or change in vocal-tract bend radius was small. A much larger formant shift was observed with changes in the volume of the piriform sinuses as well as small change in the length of the epilarynx tube. The changes in these and other local structures merit further investigation regarding their contribution to non-linear scaling of formants during childhood development.

#### ACKNOWLEDGMENTS

This work was supported in part by NIH Research Grant Nos. R03 DC4362 (Anatomic Development of the Vocal Tract: MRI Procedures) and R01 DC6282 (MRI and CT Studies of the Developing Vocal Tract), from the National Institute of Deafness and other Communicative Disorders (NIDCD). Also, by a core Grant No. P-30 HD03352 to the Waisman Center from the National Institute of Child Health and Human Development (NICHD). We thank Reid Durtschi for his assistance with anatomic measurements from the imaging studies. S.Y. is a former graduate student of University of Wisconsin-Madison.

- <sup>1</sup>L. J. Boë, J. L. Heim, K. Honda, and S. Maeda, "The potential Neandertal vowel space was as large as that of modern humans," *J. Phonetics* **30**, 465–484 (2002).
- <sup>2</sup>L. Menard, J. L. Schwartz, and L. J. Boe, "Role of vocal tract morphology in speech development: Perceptual targets and sensorimotor maps for synthesized French vowels from birth to adulthood," *J. Speech Lang. Hear. Res.* **47**, 1059–1080 (2004).
- <sup>3</sup>L. Menard, J. L. Schwartz, and L. J. Boe, "Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood," *J. Acoust. Soc. Am.* **111**, 1892–1905 (2002).
- <sup>4</sup>J. Serkhane, J. L. Schwartz, and P. Bessière, "Building a talking baby robot," *Interaction Studies* **6**, 253–286 (2005).
- <sup>5</sup>R. Harshman, P. Ladefoged, and L. Goldstein, "Factor analysis of tongue shapes," *J. Acoust. Soc. Am.* **62**, 693–713 (1977).
- <sup>6</sup>S. Maeda, "Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal tract shapes using an articulatory model," in *Speech Production and Speech Modelling*, edited by W. J. Hardcastle and A. Marchal (Kluwer, Netherlands, 1990), pp. 131–149.
- <sup>7</sup>J. Liljencrants, "Fourier series description of the tongue profile," *Speech Transmission Laboratory-Quarterly Progress Status Reports*, Vol. **12**, No. 4, Royal Institute of Technology (KTH), Stockholm, 1971, pp. 9–18.
- <sup>8</sup>P. M. C. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1987).
- <sup>9</sup>J. Sundberg, B. Lindblom, and J. Liljencrants, "Formant frequency estimates for abruptly changing area functions: A comparison between calculations and measurements," *J. Acoust. Soc. Am.* **91**, 3478–3482 (1992).
- <sup>10</sup>V. L. Streeter, *Fluid Dynamics* (McGraw-Hill, New York, 1948).
- <sup>11</sup>D. H. Keefe and A. H. Benade, "Wave propagation in strongly curved ducts," *J. Acoust. Soc. Am.* **74**, 320–332 (1983).
- <sup>12</sup>C. J. Nederveen, "Influence of a toroidal bend on wind instrument tuning," *J. Acoust. Soc. Am.* **104**, 1616–1626 (1998).
- <sup>13</sup>K. Motoki, N. Miki, and N. Nagai, "Measurement of sound-pressure distribution in replicas of the oral cavity," *J. Acoust. Soc. Am.* **92**, 2577–2585 (1992).
- <sup>14</sup>H. Matsuzaki, N. Miki, and Y. Ogawa, "3 D finite element analysis of Japanese vowels in elliptic sound tube model," *Electron. Commun. Jpn.* **83**, 43–51 (2000).
- <sup>15</sup>X. Zhou, C. Y. Espy-Wilson, S. Boyce, M. Tiede, C. Holland, and A. Choe, "A magnetic resonance imaging-based articulatory and acoustic

- study of “retroflex” and “bunched” American English /r/,” *J. Acoust. Soc. Am.* **123**, 4466–4481 (2008).
- <sup>16</sup>D. L. Brown, G. S. Chesshire, W. D. Henshaw, and D. J. Quinlan, “Overture: An object-oriented software system for solving partial differential equations in serial and parallel environments,” in Conference: 8. SIAM Conference on Parallel Processing for Scientific Computing, Minneapolis, MN (1997), pp. 14–17.
- <sup>17</sup>W. M. Chan and J. L. Steger, “Enhancements of a three-dimensional hyperbolic grid generation scheme,” *Appl. Math. Comput.* **51**, 181–205 (1992).
- <sup>18</sup>W. Henshaw, “The overture hyperbolic grid generator user guide, Version 1.0,” Research Report No. UCRL-MA-134240, Lawrence Livermore National Laboratory, Livermore, CA, 2003.
- <sup>19</sup>N. A. Petersson, “User’s guide to Xcog version 2.0,” Technical Report No. CHA/NAV/R-97/0048, Chalmers University of Tech., Gothenburg, Sweden, 1997.
- <sup>20</sup>W. M. Chan, “Hyperbolic methods for surface and field grid generation,” *Handbook of Grid Generation* (CRC, Boca Raton, FL, 1999).
- <sup>21</sup>K. Mochizuki and T. Nakai, “Estimation of area function from 3-D magnetic resonance images of vocal tract using finite element method,” *Acoust. Sci. & Tech.* **28**, 346–348 (2007).
- <sup>22</sup>M. M. Sondhi, “Resonances of a bent vocal tract,” *J. Acoust. Soc. Am.* **79**, 1113–1116 (1986).
- <sup>23</sup>H. K. Vorperian, R. D. Kent, M. J. Lindstrom, C. M. Kalina, L. R. Gentry, and B. S. Yandell, “Development of vocal tract length during early childhood: A magnetic resonance imaging study,” *J. Acoust. Soc. Am.* **117**, 338–350 (2005).
- <sup>24</sup>P. Perrier, L. J. Boe, and R. Sock, “Vocal tract area function estimation from midsagittal dimensions with CT scans and a vocal tract cast: Modeling the transition with two sets of coefficients,” *J. Speech Hear. Res.* **35**, 53–67 (1992).
- <sup>25</sup>S. Maeda, “On the conversion of vocal tract x-ray data into formant frequencies,” Bell Laboratories, Murray Hill, NJ, 1972.
- <sup>26</sup>D. Beautemps, P. Badin, and R. Laboissière, “Deriving vocal-tract area functions from midsagittal profiles and formant frequencies: A new model for vowels and fricative consonants based on experimental data,” *Speech Commun.* **16**, 27–47 (1995).
- <sup>27</sup>V. Milenkovic, “Computer synthesis of continuous path robot motion,” in *Proceedings of the Fifth World Congress Theory of Machines and Mechanisms* (ASME, New York, 1979), pp. 1332–1335.
- <sup>28</sup>M. Loo and V. Milenkovic, “Multicircular curvilinear robot path generation,” in *Robots 11 Conference Proceedings and 17th International Symposium Industrial Robots* (SME, Dearborn, MI, 1987), Vol. **18**, pp. 19–27.
- <sup>29</sup>M. Loo, Y. A. Hamidieh, and V. Milenkovic, “Generic path control for robot applications,” in *Robots 14 Conference Proceedings* (SME, Dearborn, MI, 1990), Vol. **10**, pp. 49–64.
- <sup>30</sup>W. Rostafinski, “Monograph on propagation of sound waves in curved ducts,” NASA Reference Publication No. 1248 (1991).
- <sup>31</sup>K. H. Huebner, *The Finite Element Method for Engineers* (Wiley-Interscience, New York, 2001).
- <sup>32</sup>K. Motoki, “Three-dimensional acoustic field in vocal-tract,” *Acoust. Sci. & Tech.* **23**, 207–212 (2002).
- <sup>33</sup>B. S. Atal, J. J. Chang, M. V. Mathews, and J. W. Tukey, “Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique,” *J. Acoust. Soc. Am.* **63**, 1535–1553 (1978).
- <sup>34</sup>L. L. Beranek, *Acoustics* (McGraw-Hill, New York, 1954).
- <sup>35</sup>B. H. Story, I. R. Titze, and E. A. Hoffman, “Vocal tract area functions from magnetic resonance imaging,” *J. Acoust. Soc. Am.* **100**, 537–554 (1996).
- <sup>36</sup>J. Dang and K. Honda, “Acoustic characteristics of the piriform fossa in models and humans,” *J. Acoust. Soc. Am.* **101**, 456–465 (1997).
- <sup>37</sup>T. Baer, J. C. Gore, L. C. Gracco, and P. W. Nye, “Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels,” *J. Acoust. Soc. Am.* **90**, 799–828 (1991).
- <sup>38</sup>A. Soquet, V. Lecuit, T. Metens, and D. Demolin, “Mid-sagittal cut to area function transformations: Direct measurements of mid-sagittal distance and area with MRI,” *Speech Commun.* **36**, 169–180 (2002).