THREE-DIMENSIONAL COARTICULATORY STRATEGIES OF TONGUE MOVEMENT

M. Stone[1], A. Lundberg[2], E. Davis[2], R. Gullapalli[3], and M. NessAiver[3]
[1]University of Maryland School of Medicine
Division of Otolaryngology
16 S. Eutaw Street, Room 525, Baltimore, MD 21201
Tel. (410) 328-1888, FAX: (410) 328-5827, E-mail: mstone@surgery2.ab.umd.edu
[2]Johns Hopkins University
[3]University of Maryland Medical Systems

ABSTRACT

This paper will present three-dimensional tongue "volumes," reconstructed from three sagittal slices (left, mid, right) made using tagged cine MRI. The volumes will be animated to show CV movement from the consonants /k/ and /s/ to the vowels /i/, /a/, and /u/.

INTRODUCTION

The tongue is a three-dimensional non-rigid body, which is normally measured as a two-dimensional object using images or a series of points in the midsagittal plane. Three-dimensional static data are available, but there is no 3D motion of the tongue available. This paper presents three-dimensional tongue "volumes," which move in time through up to seven time-phases. The volumes are reconstructed from three sagittal slice-series (left, mid, right) made using tagged cine MRI. Each slice-series is captured at 7 time-phases during the course of one second. The series are animated to show movement for six Consonant-Vowel (CV) signatures. As the tongue deforms over time, the measured points reflect both surface and internal deformations.

The project explores the coarticulated movements from a front or back consonant (/s, k/) to the three point vowels (/i, a, u/). Our previous research has depicted 3D static tongue surface shapes for 18 English sounds [1]. Based on that, the two consonants /k/ and /s/ were chosen for this study. The chosen sounds differ in static shape as well as position. The /s/ is a front consonant with a continuously grooved tongue body from front to back. The /k/ is a back consonant with an arched tongue body from front to back. We were interested in examining the 3D strategies of motion from the arched /k/ to the slightly grooved /a/, to the front raised /i/, which has an arched anterior and grooved posterior shape, and to the /u/, whose shape is not too different from the /k/ itself. The movement from the grooved /s/ into those three vowels was expected to be quite different. The differences should provide insight into the tongue's strategies of three-dimensional surface and internal motion.

While one can predict the nature of tongue shape changes using simple linear interpolation between the extreme positions, it is well established that tongue motion over time is not linear. As early as 1973, tongue movement was modeled as a cosine function [2]. Tongue movements in the sagittal and coronal planes have indicated that different portions of the tongue surface reach the maximum phoneme target at different times and do not move toward it simultaneously [3]. The present data cannot track precisely the time course of the CV motion, because the sampling rate was only 7 Hz. However, the path of local displacements, and the component deformities that contribute to the global shape change are presented below.

Questions asked by this study are: (1) Are the effects of consonants on vowels greater than that of vowels on consonants? C and V endpoint shapes for the three planes are compared to determine the most and least consistent shapes. (2) Do local portions of the tongue move linearly in time from target to target? Pathlines are plotted to reveal the direction and rate of movement of points on and within the tongue's surface in each plane. (3) How is lateral tongue movement different from midline? X-shift, y-shift, and xy-shear are calculated to compare the C-to-V changes in all three planes.

METHODS

A 19-year-old male, native speaker of English repeated the CV's /si, sa, su, ki, ka, ku/ 96 times each while a left, mid, and right sagittal cine-MRI slice-sequence was collected. The three MRI slices are 7 mm thick and 4 mm apart in space. Each slice was encoded with a grid of tag lines during production of the consonant, and seven time phases were recorded during the succeeding second as the tongue moved into the vowel. Each CV was repeated 32 times to create a single slice sequence (96 times for the three-slice series). The tags are made of intersecting horizontal and vertical lines each 1.3 cm apart. As the tongue moves between two sounds, the tags deform, reflecting the internal tissue deformation. The tags are fairly dense, and as the tongue moves quickly, the deformations become quite large (with respect to the tag grid spacing). For fast tongue movements, additional time steps were interpolated between the collected temporal data sets to facilitate tag alignment.

The tag line intersections and the points along the tongue's circumference were stored as xy coordinate points. The left, middle, and right sets of points at each time phase were aligned in 3D space using Vutongue, custom software discussed in Stone and Lundberg (1996). Z coordinates were set so that the interslice distance was twice that of the original distance, making the midline slices more visible. The 3D time series of tag points and slice edges (real data and interpolates) were then imported into a 3D visualizing program (Geomview, Center for Geometric Computing) for interactive display and export to videotape.

The deformations observed between slices were decomposed using a continuum mechanics optimization method. The deformation components used were x- and y-stretch, xy shear, rotation, and translation. This compact deformation set was designed to illuminate the coarticulatory effects of C-to-V movement by breaking down the total deformation into quantitatively discriminated individual effects. The process employed involved sampling the edge of the tongue in both the reference (C) and the deformed (V) state and then running custom continuum mechanics software to optimally separate the components of the deformation based on the change of external shape [4]. Internal movement can be inferred from the results and compared to the observed deformations of the tag lines. The x- and y-stretch values in Table 1 represent percentage change from the original shape. The xy-shear values, when squared, are comparable in magnitude to the stretch values.

RESULTS

One of the features of interest studied in this project was the variability of phoneme shapes across contexts. Figure 1 shows overlaid surfaces of /k/ and of /s/ in the three vowel contexts, and for /i/, /a/, and /u/ in the two consonant contexts. Visual inspection indicates the most overlap for /u/, progressing to the least overlap through /k/ /i/ /a/ /s/.

Table 1. Surface deformations for each CV

	x-stretch	**y-stretch**	**xy-shear**	**x trans**	**y trans**
LEFT	.06	-.14	.0081	-.08	-.06
k to a MID	.13	-.21	.0025	-.08	-.08
RIGHT	.15	-.13	.000064	-.08	-.11
LEFT	-.09	-.05	.0009	.03	-.05
k to u MID	-.03	-.02	.000064	.01	-.01
RIGHT	-.02	-.05	.0025	.006	-.01
LEFT	-.04	-.08	0	-.02	-.07
k to I MID	-.02	-.03	.0004	.006	-.009
RIGHT	.07	.05	.0001	-.03	.001
LEFT	.003	.009	.000004	-.02	.007
s to a MID	.06	-.01	.0289	.01	.03
RIGHT	.06	-.009	.0004	-.03	.01
LEFT	-.02	-.03	.0009	.01	.04
s to u MID	-.05	.04	.0001	.01	-.01
RIGHT	-.06	.17	.0016	.02	.08
LEFT	.001	-.05	.0001	.007	.007
s to I MID	-.05	.01	.01	.02	.09
RIGHT	-.01	.002	.0049	.06	-.05

A second area of interest was global changes in tongue surface shape in the midline versus lateral planes. Table 1 shows five component deformations that characterize the change in surface shape between the consonant and vowel. The tags were not used in these calculations. The x- and y-stretch values are percentages, with positive numbers referring to elongation. Xy-shear indicates stretch upward and backward (positive) or upward and forward (negative). The shear values in the table are squared to make them comparable to the stretch measures with respect to their effect on area. In the present data set, the shear values are all very small, and the CV movements themselves tend to be small. The largest deformations occurred for /ka/ whose surfaces are shown in Figure 2. At midline and on the right, the tongue compresses in the y-dimension and expands in the x-dimension, which is reflected in Table 1. On the left, only y-compression is observed. An interesting asymmetry is observed for /su/, where y-expansion occurs only on the right, reflecting rotational movement as the tongue moves upward from /s/ to /u/ on the right primarily. Pathlines could not be calculated for /su/ because the tags were laid down too late in the sequence.

The third area examined in this paper is local changes in direction and rate during the CV movement. Linear pathlines were drawn connecting the points in each frame. Figure 3 presents pathlines for mid and lateral movements of /ku/. Each line contains an arrow after the third and fifth frames indicating direction of movement. It is immediately apparent that the CV movement is non-linear. Although the overall motion is downward, other directions are represented as well. On the left, Table 2 indicates an x-compression in the surface shape. This is visible in Figure 3A and the internal tags reflect this posteriorward movement toward the end. At midline, the tags reveal that the upper tongue lowers more rapidly and to a greater extent than the rest, and in all three slices the tags in the tip compress backward.

DISCUSSION

These results are an interesting way to examine coarticulation. In the static shapes, neither C's nor V's were more variable. Rather the high sounds /u/ and /k/, which were plastered against the palate, varied less.

The pathlines revealed that the largest movements occurred at midline in the upper 1/3 of the tongue, and posteriorly for /i/. It did appear that the movement patterns were sensitive to the constrictions. The /ka/ and /ku/ movements were primarily downward. For /sa/, movement was down and back. At midline and left, /ki/ was downward, on the right, there was a little upward movement. Movement for /si/ was up and back at midline, downward on the left, and forward on the right. For the high sounds, the top had greater movement.

Global deformations served several purposes. They reflected differences in magnitude of contribution of each deformation to the whole shape change. They also revealed asymmetrical movements, such as for /si/ (x-expansion on right) and for /su/ (y-expansion and positive y-translation on right).

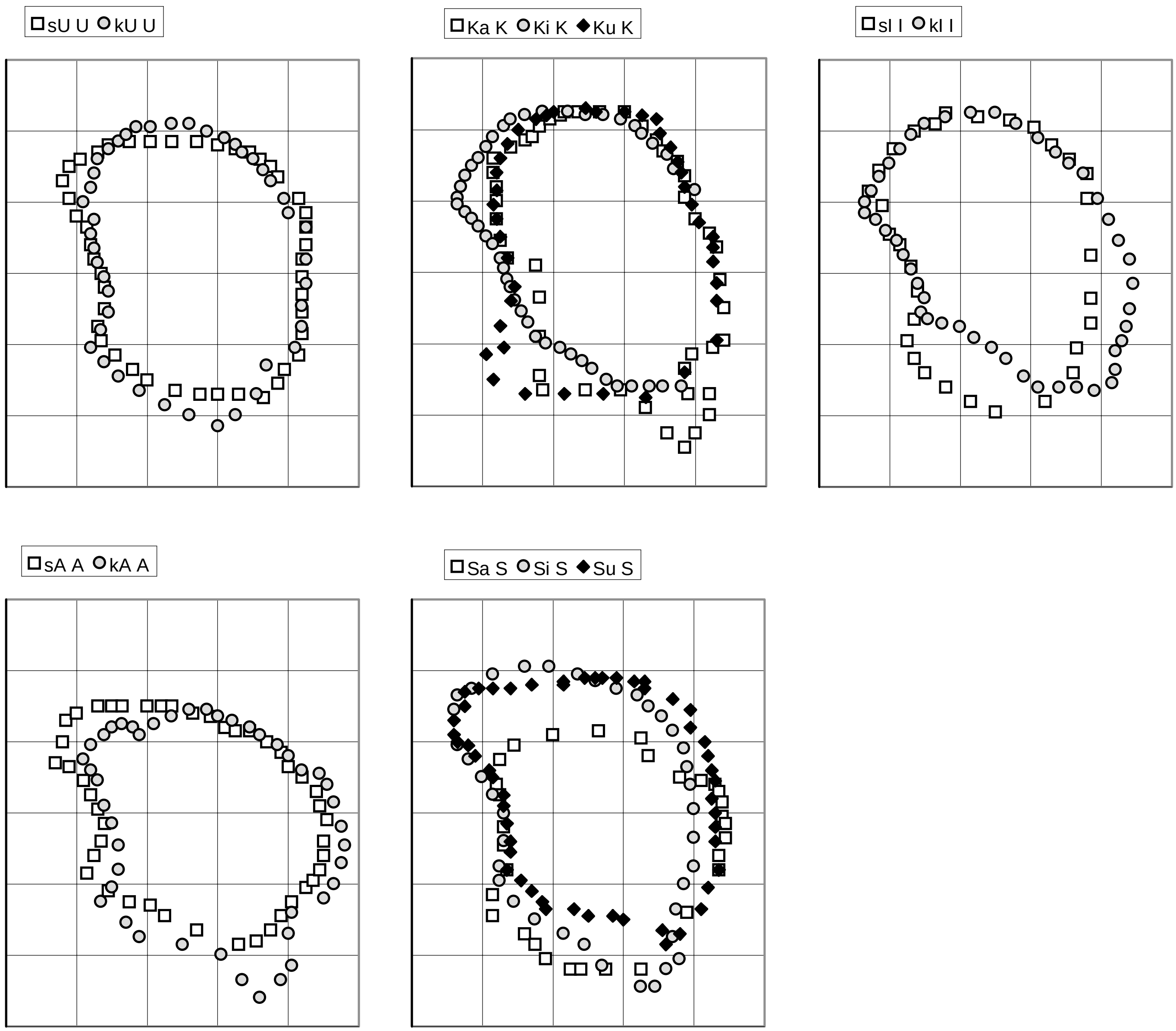

Figure 1 a-e. Overlaid drawings of each sound taken from the different contexts.

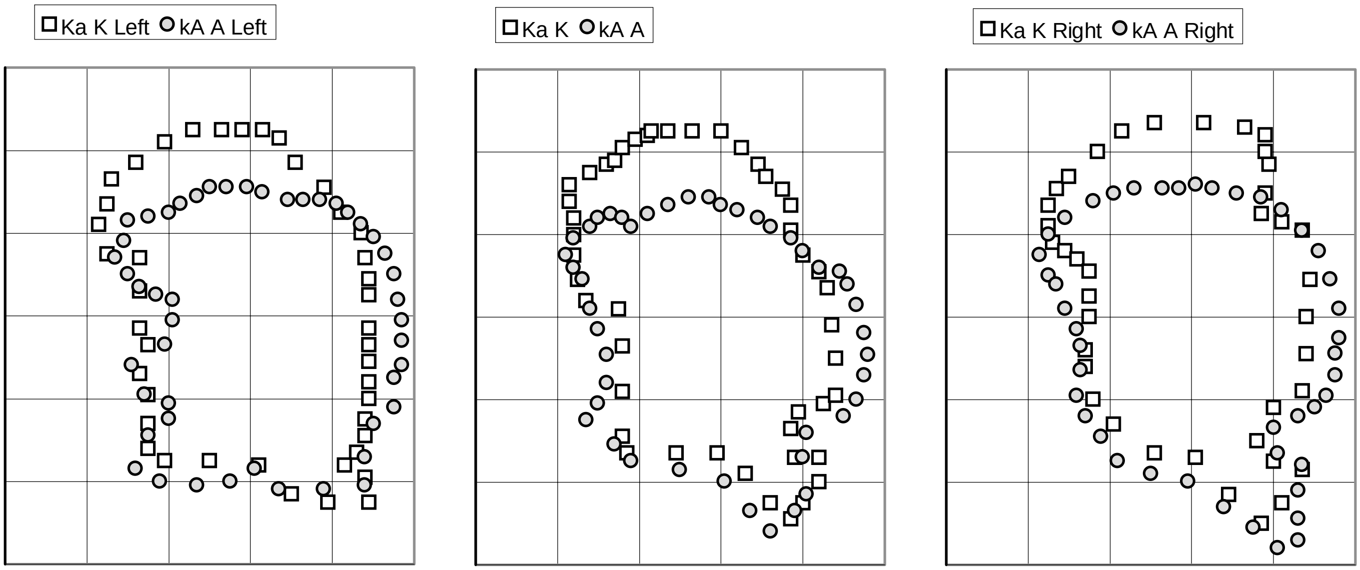

Figure 2 a-c. Surface movement from /k/ to /a/.

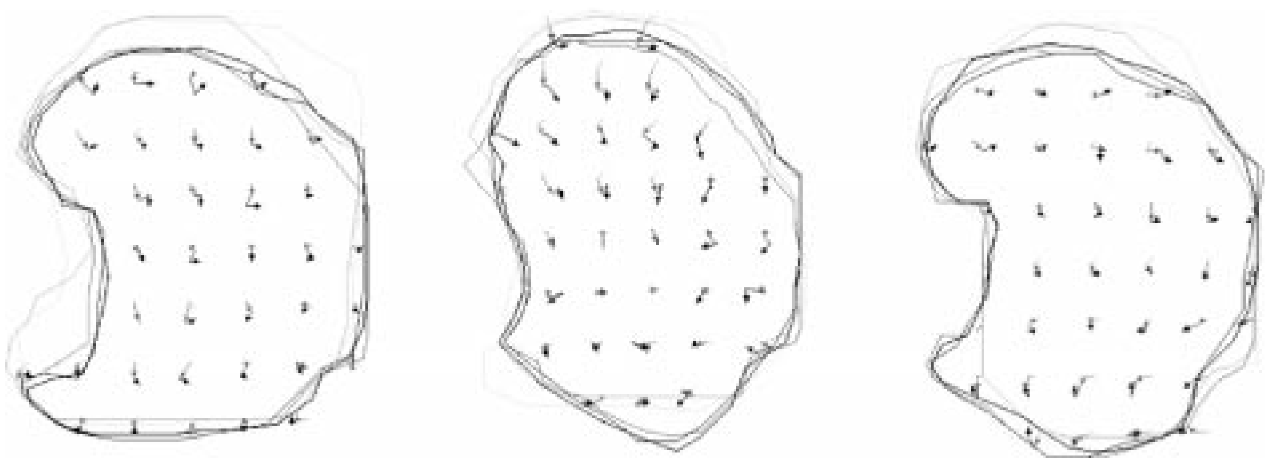

Figure 3 a-c. Pathlines indicating direction and extent of tag movement over 5 frames from /k/ to /u/.

REFERENCES

[1] M. Stone and A. Lundberg, "Three-Dimensional Tongue Surface Shapes of English Consonants and Vowels", *J. Acoust. Soc. Amer.*, Vol. 99, pp. 3728-3737, 1996.

[2] P. Mermelstein, "Articulatory Model for the Study of Speech Production", *J. Acoust. Soc. Amer.*, Vol. 53, pp. 1070-1080, 1973.

[3] M. Stone, "A Three-Dimensional Model of Tongue Movement", *J. Acoust. Soc. Amer.*, Vol. 87, pp. 2207-2217, 1990.

[4] E. Davis, A. Douglas, and M. Stone, "*A Continuum Mechanics Representation of Tongue Motion in Speech*", Proceedings of the 4th International Conference on Spoken Language Processing, Philadelphia, PA, 1996.