JSLHR

Research Article

Tongue Motion Patterns in Post-Glossectomy and Typical Speakers: A Principal Components Analysis

Maureen Stone,[a] Julie M. Langguth,[a] Jonghye Woo,[a] Hegang Chen,[b] and Jerry L. Prince[c]

Purpose: In this study, the authors examined changes in tongue motion caused by glossectomy surgery. A speech task that involved subtle changes in tongue-tip positioning (the motion from /i/ to /s/) was measured. The hypothesis was that patients would have limited motion on the tumor (resected) side and would compensate with greater motion on the nontumor side in order to elevate the tongue tip and blade for /s/.

Method: Velocity fields were extracted from tagged magnetic resonance images in the left, middle, and right tongue of 3 patients and 10 controls. Principal components (PCs) analysis quantified motion differences and distinguished between the subject groups.

Results: PCs 1 and 2 represented variance in (a) size and independence of the tongue tip, and (b) direction of motion of the tip, body, or both. Patients and controls were correctly separated by a small number of PCs.

Conclusions: Motion of the tumor slice was different between patients and controls, but the nontumor side of the patients' tongues did not show excessive or adaptive motion. Both groups contained apical and laminal /s/ users, and 1 patient created apical /s/ in a highly unusual manner.

Key Words: glossectomy, MRI, tongue motion, principal components, speech

The National Cancer Institute estimated that 42,000 people in the United States would develop oral or oropharyngeal cancer in 2014 and that 32% of them would be in the tongue (American Cancer Society, 2014). The most prevalent treatment is a *glossectomy,* or surgical removal, of all or part of the tongue. Postoperative defects in tongue size and form as a result of glossectomy can have detrimental effects on patients' speech, which in severe cases impact patients' intelligibility (cf. Nicoletti et al., 2004). Loss of intelligibility reduces communicative ability and can result in reticence to speak in social situations or in avoiding them altogether, resulting in a potentially significant public health problem.

Numerous variables are possible determinants of postoperative tongue motility. Loss of the tongue tip is more detrimental to speech than is lateral tongue removal with tip preservation (Logemann et al., 1993; Michiwaki, Schmelzeisen, Hacki, & Michi, 1993). This is not an issue in the present study. All the patients who participated in this study had a preserved tongue tip, and a tumor on the lateral border of the tongue behind the tip. Size of lesion is another key parameter affecting speech. Pauloski, Logemann, Colangelo, Rademaker, and Esclamado (1998) found a 30% correlation between the percent of oral tongue resected and the quality of /s/. Nicoletti et al. (2004) found a statistically significant correlation between size of resection and intelligibility. The present study controls for this by using patients with tumors of roughly the same size. The preservation of midsagittal motility is also important to speech, and especially /s/, which requires a midsagittal groove throughout its length. Bressmann, Ackloo, Heng, and Irish (2007) studied concavity and symmetry in pre- and postsurgery glossectomy patients and controls. Concavity was significantly reduced postoperatively for flap-closure patients but not for primary-closure patients. In addition, a moderate positive correlation was found between postoperative speech acceptability ratings and the degree of change in concavity. No correlation was found for changes in left-to-right symmetry. The three slices studied in the present work should reveal motion differences between the midline and lateral slices, or the left and right sides. Rastadmehr, Bressman, Smyth, and Irish (2008) studied rate of motion in the anterior, central, and posterior portions of the midsagittal tongue for changes

[a]University of Maryland School of Dentistry, Baltimore
[b]University of Maryland School of Medicine, Baltimore
[c]Johns Hopkins University, Baltimore, MD

Correspondence to Maureen Stone: mstone@umaryland.edu

Editor: Jody Kreiman
Associate Editor: Caryn Easterling

Received April 4, 2013
Revision received June 19, 2013
Accepted July 19, 2013
DOI: 10.1044/1092-4388(2013/13-0085)

__Disclosure:__ The authors have declared that no competing interests existed at the time of publication.

in motion velocity. They found that before glossectomy surgery, patients moved their posterior tongues significantly more slowly than controls; after surgery, the same patients moved the center part of their tongue faster than controls. The principal components in the present study should reveal any differences in local rate among the subject groups.

The /s/ sound is produced in one of two ways by typical speakers (cf. Dart, 1998). For an *apical* /s/, the tongue tip elevates to form a constriction at the alveolar ridge and direct the airstream toward the incisors. For a *laminal* /s/, the tongue blade elevates to make the constriction and the tongue tip is lower. Post-glossectomy speakers, who have lost a lateral section of the tongue body, lose some innervation of the tongue tip, even when the tip is preserved. We expect that the apical versus laminal /s/ will be seen as one of the /s/ features captured by the PCs, and propose that patients will be more likely to use the laminal /s/ because it does not require the subtle, independent elevation of the tip required by the apical /s/.

The goal of the present study is to quantitatively represent motion patterns internal to the tongue as a way to determine the effects of anatomical changes due to surgery on tongue behavior during speech. To better understand the effects of the surgery throughout the tongue, three parallel sagittal slices were measured: one slice each through the tumor side, the nontumor side, and the midline in patients. In controls, both sides were measured. The expectation was that for patients, the nontumor side would compensate for rigidity on the tumor side to achieve a typical midline motion. In addition, both sides were expected to exhibit unusual motion patterns relative to the controls, which might only be captured in the higher order PCs.

Magnetic resonance imaging (MRI) creates images of the nuclear magnetic resonance properties of large collections of hydrogen atom nuclei (Bushong, 2003). Because the tongue contains hydrogen in both water and fat in spatially varying proportions, MR images of the stationary tongue show details of its muscle anatomy. However, it is difficult to capture tongue motion—during speech, for example—because shorter imaging times yield much lower signal levels while noise remains constant. Faster imaging techniques permit the capture of motion, but these images are degraded both in resolution and tissue contrast. To improve image quality in order to permit motion visualization and analysis, subjects are trained to repeat a word multiple times to the timing of a metronome. Different (Fourier) components of the image data are acquired with each repetition and are combined to form a single, high-quality image sequence—so-called *cine-MRI*—which depicts the tongue's movement during a single repetition.

Because of the need for fast acquisition, tissue contrast in cine-MRI is low, and it is difficult to discern the internal (muscle) details of tongue motion. *Tagged-MRI,* originally developed to measure heart motion (Axel & Dougherty, 1989), can be used to create temporary patterns in the tongue muscle, from which motion can be measured. Tagged-MRI works by magnetically marking, or tagging, planes of tissue prior to collecting the MR image sequence. When the MRI sequence is collected, the motions of the tagged planes are visible in each time-frame. When measuring 2-D tongue motions, the midsagittal plane of the tongue is an excellent slice to use because it is, perhaps, the best single representative of the total motion. This is because most sounds use bilateral contact between the tongue and palate or inner tooth surfaces to create a midsagittal air channel (Stone & Lundberg, 1996). A left and right plane are also critical for this study because the patients are expected to show asymmetries not found in the controls.

Principal components analysis (PCA) is a statistical method that reduces dimensionality of a data set to represent complex patterns of variance using its major components. Velocity fields are high-dimensional data sets, with hundreds of tissue points moving in complex patterns. PCA reduces data dimensionality by finding uncorrelated variables, called *principal components (PCs),* which explain the variance in the observed data in descending order from highest to lowest variance starting with the first PC. The directions of the PCs represent orthogonal variations in the data that can be visualized in conjunction with our understanding of the physical nature of the measurements, and are often behaviorally interpretable. PCA has been used previously to capture patterns in velocity fields of the tongue (cf. Stone, Liu, Chen, & Prince, 2010). PCA allows one to examine and quantify subtle feature differences among subjects and to visualize those differences. The fewer the PCs needed to distinguish two groups with 100% accuracy, the more the groups differ.

The patients chosen for this study all had quite similar resection properties. They had a tongue tip on one side with reduced innervation and a break in the muscle and nerve tissue leading to it. Their lateral resection was closed primarily and restricted motion because tissue that had been previously farther apart was now sewn together. In addition, the healing process postglossectomy leaves a scar that is rigid and cannot stretch as easily as the adjacent tissue. The small elevation of the tongue tip from /i/ into /s/ is accompanied by a large cross-sectional shape change from a flat/arched surface to a grooved surface. The overall motion is small, but the 3-D coordination needed is large. Therefore, we proposed three hypotheses regarding how patients adapt to these changes: (a) Patient motion patterns will be smaller than is typical on the tumor side of the tongue and contain unique or unusual motion features. (b) Patients' tongue motion on the nontumor side will be greater than is typical because it is compensating for the tumor side. (c) Patients will more likely use a laminal /s/, which requires less elevation of the tongue tip and will be easier for those with reduced control of the tip.

Method

Subjects and Speech Materials

Subjects included 10 typical controls and three glossectomy patients seen at least 7 months postsurgery (see

Table 1). All subjects were native American English speakers. Informed consent and all procedures were approved by the Institutional Review Board at the University of Maryland, Baltimore. Subject ages were M_{age} = 27.3 years (SD = 8.9) for the controls and M_{age} = 44.3 years (SD = 15.5) for the patients. The large difference in age is due to the early stages of the research and the difficulties in finding matched controls who tolerated MRI and had few fillings. Subjects 11–13 were the patients. Subject 13 was missing multiple maxillary left and right posterior teeth and had a torus in the palatal vault. All other subjects had complete dentition and no palatal anomalies. Subject 13 also was scanned only 7 months postsurgery and may not have recovered completely from the surgery. The glossectomy patients repeated unique versions of the Sentence Intelligibility Test (SIT; Yorkston & Beukelman, 1981) and were rated by a speech-language pathologist as having highly intelligible speech (see Table 1). None of the patients had speech therapy.

The tumors were staged according to the tumor-node-metastasis (TNM) classification system, in which T refers to tumor size, N to lymph node involvement, and M to metastasis (Greene et al., 2002). The three patients all had a squamous cell carcinoma in the left or right lateral margin of the tongue, and in the medial third of the tongue, not the anterior or posterior third. The tumor and 1 cm–1.5 cm of healthy tissue were removed by a partial glossectomy (see Table 1). The tumors were all T1N0M0 (*T1* = less than 2 cm in the greatest diameter; *N0* = no clinically positive lymph nodes; *M0* = no evidence of metastasis). All patients received primary closure after excision of the malignancy—that is, the wound was closed by sutures. The patients were not scanned with MRI presurgically for several reasons: Their surgery was scheduled as expeditiously as possible, limiting the time available; they were often in pain and would have been uncomfortable performing the study; and, most important, tumor bulk and discomfort cause patients to modify their tongue motion even in the presurgical condition. Thus, MRI data may not be representative of their typical speech. Acoustic analyses of fricatives has shown that presurgical patients have atypical /s/ spectra (Zhou, Stone, & Espy-Wilson, 2011).

The MRI speech material was the phrase *a geese*. This phrase was chosen, despite its poor grammar, for several reasons. Phonetically, the task begins with a neutral vocal tract configuration (schwa), the tongue body motion is fairly simple as it moves only anteriorly, and the word uses little to no jaw motion, increasing the potential for tongue deformation. The speech sample length was necessitated by MRI restrictions: MRI tags fade in 1.2 s; therefore, speech data are collected for only 1 s to ensure high-quality tag data throughout the sequence. Data for other words were collected, but this task was studied here because of its phonemic content. The motion between /i/ and /s/ might elicit a different transition mechanism from what is typical because /s/ is a challenging sound, and the transition from /i/ to /s/ requires tongue tip elevation and lowering of the entire midline tongue into a groove. These might add additional difficulty to speakers with nonsymmetric tongue musculature.

MRI Instrumentation and Data Collection

All MRI scanning was performed on a Siemens 3.0 T Tim Trio system using an eight-channel head and neck coil. Sagittal "stacks" of data were collected using two MRI methods: cine-MRI and tagged-MRI. The cine and tagged stacks had the same imaging parameters: identical slice locations, 6-mm slice thickness, 6-mm tag spacing, 1.875 mm × 1.875 mm in-plane spatial resolution, 1-s recording time, and 26 time-frames per second. Depending on tongue size, either seven or nine sagittal slices were collected of cine- and tagged-MRI data. Both MRI methods produce a single "movie" for each slice by acquiring multiple repetitions and postprocessing the data using ensemble summation to add all comparable time-frames across the multiple repetitions of the speech task.

Tagged-MRI data were collected using complementary spatial modulation of magnetization (CSPAMM) tagging (Fischer, McKinnon, Maier, & Boesiger, 1993) and reconstructed using magnitude image CSPAMM reconstruction (MICSR; NessAiver & Prince, 2003). CSPAMM acquisition first acquires a cosine tag pattern and then a minus cosine tag pattern. Standard CSPAMM reconstruction subtracts these two images yielding a perfect cosine tag pattern (distorted by any motion that may occur) regardless of how much the tags fade (Fischer et al., 1993). MICSR uses the same two acquisitions but combines them using only the magnitude data (without requiring their phase). In addition to avoiding the need for collection of phase data, MICSR has improved contrast-to-noise ratios over CSPAMM combination, especially at later times in the image sequences (NessAiver & Prince, 2003).

Table 1. Individual patient data.

	Patients		
Variable	**Subject 11**	**Subject 12**	**Subject 13**
Resected side	right	right	left
Resection size (cm)	3.5 × 2.5 × 2.2	2.4 × 2.1 × 1.8	3.5 × 1.9 × 2.2
Time post-op	39 mo.	46 mo.	7 mo.
Sentence Intelligibility Test	100	99	96
Acceptability (0 = *best,* 10 = *worst*)	0.25	0.7	2.8
/s/ type	apical	laminal	apical

An MICSR data set is composed of four data acquisitions. Two of them contain horizontal tags and two contain vertical tags; each tag direction is acquired twice, one with a cosine tag pattern and one with a minus cosine tag pattern. Each of these four acquisitions requires three repetitions per slice, in order to acquire adequate Fourier data for analysis. Thus for seven sagittal slices there are 21 repetitions of the task collected as four separate acquisitions, with three intervening pauses (Parthasarathy, Prince, Stone, Murano, & NessAiver, 2007). The cine-MRI data set requires five repetitions per slice (in order to acquire adequate Fourier data), resulting in 35 repetitions for seven slices. The entire cine-MRI and tagged-MRI data collection takes 20 min per word.

Speaker precision was optimized by training the subject to speak to the same metronome beat that is used in the scanner. Each tagged-MR image sequence is a combination of multiple repetitions. Therefore, subject variability across repetitions will cause blurring when the images are combined. To maximize speaker precision, subjects were trained prior to the MRI scan to precisely repeat the speech tasks. Subjects were also trained to inhale and exhale at fixed places within each cycle to further improve task repetition. The training used a metronome with a four-beat sequence (for two syllables, inhalation, exhalation) based on the work of Masaki et al. (1999).

Speech recordings were made in the MRI scanner with a subtraction-type fiber optic microphone (Optoacoustics Ltd., Moshav Mazor, Israel) with no metallic components. These data were used only to corroborate accuracy of phoneme segment breaks (Boersma & Weenink, 2010) and are not discussed further.

Data Analysis

Usually, motion of the tagged-MRI planes is tracked and motion of the points between the tags is interpolated. The present study uses harmonic phase (HARP) imaging and image processing methods, which were developed to track every tissue point in the tongue independently with no interpolation, resulting in more reliable tracking of tissue motion (Osman, Kerwin, McVeigh, & Prince, 1999). HARP tracks phase relationships in the tag deformations, across all 26 time-frames, to estimate the motion of every pixel in the image (Parthasarathy et al., 2007). From these pixel motions, running velocity fields within the tongue were calculated between adjacent frames, in the midsagittal slice, and in a right and left parasagittal slice. The parasagittal slices were separated by one slice from the midsagittal slice. Thus the distance was 12 mm from slice-center to slice-center. This allowed enough interslice distance to capture lateral tongue motion, while avoiding slices that had missing tongue tissue for the patients.

These three planes were studied by defining them as vector fields of tissue point motion. The vector field used in this study is the *velocity field,* which quantifies the direction and velocity of every tissue-point (*voxel*) in the sagittal slice. Velocity fields were calculated between each consecutive time-frame. The time-frame containing the maximum velocity between /i/ and /s/ in *a geese* was individually selected from the sequence for each subject and each slice using three steps: (a) The temporal location of /i/ and /s/ was identified in the audio recording from the MRI scanner. (b) The most rapid tongue surface motion was identified in the cine-MRI images during that acoustic transition. (c) The fastest velocity field was identified in the tagged-MRI data just at or before the fastest tongue surface motion. This velocity field became the "key time-frame" representing the /i/-to-/s/ transition in subsequent velocity analyses. The tongue portion of the key time-frame contained about 400 pixels. Their velocities were averaged to provide an average tongue velocity for each key time-frame in each slice. These averages were used only to determine the side with larger and smaller motion in controls; some sides had almost equal motion. For patients, the tumor side was always classified as the side with smaller motion and the nontumor side as the one with larger motion (see Table 1).

PCAs were performed separately and independently on the key time-frames of each of the three MRI slices: midsagittal, tumor/small motion, and nontumor/large motion. There were two possible ways to perform the analyses: The PCA could be performed on the controls and the resulting PCs applied to the patients' data, or the PCA could be performed on all 13 subjects. The former method would indicate how well the typical PCs represented the patients, but if the patient motion patterns were extremely aberrant, the control PC motion patterns would not provide insight into the patient motion patterns. The inclusion of patient data in the analysis itself would ensure that the patients' motions were represented in the PC shapes, but would not perfectly capture typical motion. To decide which method to use, two PCAs were performed on the key time-frames in the midsagittal slice, one using the 10 controls and another using all 13 subjects. The resulting PC1 and PC2 shapes were almost indistinguishable, and it was decided to include the patients in the PCA so that the PCs would include their aberrant motion components, even if only the higher PCs, that is, those with lower eigenvalues.

Before computing the PCAs, we registered landmark points for all participants, because PCA requires that all data sets be the same size and in the identical coordinate space. Registration was based on the linear alignment of nine landmark points selected just below the tongue surface for each subject (see Figure 1). Three types of points were used as landmarks: Type 1 included fairly reliable tissue points, such as (2) base of valleculae, (5) high point of the tongue, (7) tongue tip, and (8) the point below the tip where the anterior tongue surface turns from horizontal to vertical (usually the floor of the mouth in schwa). Type 2 contained points measured on specific trajectories, such as (1) the point below the valleculae just deep to the soft tissue/air interface, (4) the intersection between the tongue surface and the line drawn from the inner aspect of the mandible to the bend of the raised velum (the bend is marked with an *x* in Figure 1), and (9) the point directly inferior to Point 8 just within the soft tissue/air interface. Type 3 contained points midway between two other points, such as (3) the point

Figure 1. Nine landmark points. These points are selected from the cine–magnetic resonance imaging (MRI) scan of each subject and used to register the surface tongue shape across subjects prior to principal components analysis (PCA). The velar bend is marked with an *x*.

equidistant between Points 2 and 4, and (6) the point equidistant between Points 5 and 7. The landmark points were selected manually on the key time-frame and were as identical as possible across the 13 subjects and three slices. The points were selected on the cine-MR images because the tongue surface is not imaged well in the tagged-MR images. In addition, the points were selected just below the tongue surface because tag tracking is less accurate at the tissue–air interface.

Quality of the landmark points was examined with HARP analysis. The landmark points selected in the cine images were overlaid on the tagged data set. Two problems were possible with the data: The landmark points might be too near the surface and might mistrack, and the subject might have used slightly different tongue positions in the cine and tagged recordings, reducing the reliability of the landmark point overlay. To reduce the first possibility, each point was tracked across all 26 time-frames and checked for tracking accuracy. If mistracking occurred, a nearby point, usually deeper within the tongue, was selected instead. To reduce the second possibility, the motion of the entire word was visually compared across the two data sets to be sure the tongue motion and shape were at least visually consistent across the two data sets. However, it is not possible to completely rule out some mismatch between the two data sets and a possible difference in positioning of landmark points among subjects. The landmark points for each subject were connected linearly and registered to Subject 1 (chosen arbitrarily as the reference subject) using landmark-based registration that included rotation, translation, and scaling of the common area.

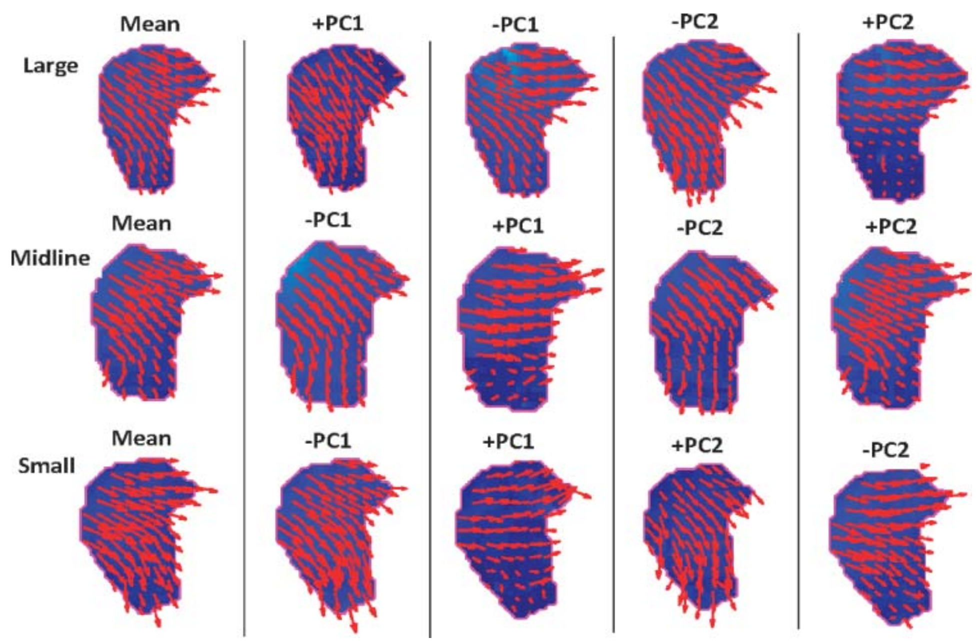
PCAs were then performed independently on the large motion/nontumor, the midline, and the small motion/tumor slices to quantify the component motions of their velocity patterns as orthogonal components (PCs) representing variance in features of motion. This was done by projecting the high-dimensional features (motion of 400+ points) in the tongue into a low-dimensional space (small number of PCs), which was used to distinguish patients from controls. The PCs were then input to *t* tests and linear discriminant analysis (LDA) to distinguish between groups. For each PC, *t* tests were used to detect the difference in loadings between controls and patients with a *p* value of .05. These were two-sided tests, with no direction specified. LDA was used to separate the patients and controls into distinct groups based on their PC representations. For this representation, we sequentially increased the number of PCs entered into the LDA, starting from PC1, PC1 + PC2, and so forth, until LDA achieved 100% accuracy. In addition, models of the underlying data set motion patterns were made by adding and subtracting a specific amount of the PCs that were statistically important. The models in Figure 2 use three dimensions, the mean, PC1, and PC2, to depict the key features of motion in the data. See Stone et al. (2010) for mathematical methodological details.

Results

PC Representation of Motion Variance in the Three Slices

Table 2 displays the percentage of variance accounted for by each of the first three PCs. The amounts were similar for all three tongue slices: PC1 represented about 40% of the variance and PC2 about 23%. The first two components were studied in depth to determine the appearance of their features, because they accounted for almost two thirds of the variance in the motion patterns. To do this, the motion patterns for the three tongue slices were modeled by adding 1 *SD* of ±PCs 1 and 2 to the mean. Figure 2, Column 1, depicts the mean velocity field for each of the three sagittal slices, large motion/nontumor, midline, and small motion/tumor (top, middle, and bottom, respectively). To the right are the models of the velocity fields that result when 1 *SD* of ±PCs 1 and 2 are added to the mean. An individual speaker's velocity fields can be represented by a model of the mean plus the combination of their loadings on any number of the PCs. If all the PCs were added to the mean, in the loading proportions of a specific speaker, that person's original velocity field would be recreated. Models composed of the mean plus the subject's first few PC loadings reflect how well the individual subject is represented by the most common directions of variance, and how well the model fits

Figure 2. PCA results. Mean velocity field for each slice (left) is followed by explanatory models composed of the mean velocity ±1 *SD* of the first two principal components (PCs).



their data. Figure 2 is an abstract representation that models the patterns that result from adding or subtracting 1 *SD* of PCs 1 and 2. Although no single subject has these exact patterns of midline motion, the figures decompose the effect of the two largest directions of variance in the data.

Let us examine the three tissue slices in Figure 2 qualitatively. The top row models the mean ±PCs 1 or 2 in the large motion/nontumor slice. The PC1 models demonstrate notable variance in tongue-tip size and direction. The PC2 models capture variance in forward versus downward motion of the tongue body and smaller versus larger motion in the tongue root. The second row contains midline mean and PC models. The PC1 models represent variance of the entire tongue from more downward to more forward. The PC2 models capture tongue-tip elevation and lowering. In the small-motion/tumor slice, reduced tip size is seen in both PC1 models. The PC1 models also reflect variance in the entire tongue from downward to forward. Variance in tongue motion direction is also represented by ±PC2, and tongue-tip independence is virtually eliminated by ±PC2. Some variance in divergence/convergence between the upper and lower tongue is represented in the mean ±PC2 models.

Table 2. Percentages of variance explained by each principal component (PC).

	Large/nontumor slice		Midsagittal slice		Small/tumor slice	
PC	**%**	**Cum. %**	**%**	**Cum. %**	**%**	**Cum. %**
PC1	41.2	41.2	42.1	42.1	38.9	38.9
PC2	23.1	64.3	22.4	64.5	23.5	62.4
PC3	9.8	74.1	9.9	74.4	11.0	73.4
PC4	7.4	81.5	5.9	80.3	6.1	79.5
PC5	5.7	87.2	5.4	85.7	5.2	84.7
PC6	4.2	91.4	3.5	89.2	4.0	88.7
PC7	2.6	94.0	2.7	91.9	3.7	92.4
PC8	2.0	96.0	2.2	94.1	2.1	94.5
PC9	1.4	97.4	1.9	96.0	1.7	96.2
PC10	1.3	98.7	1.5	97.5	1.6	97.8
PC11	0.9	99.6	1.4	98.9	1.2	99.0
PC12	0.5	100.1	1.1	100.0	1.0	100.0

Note. Cum. = cumulative.

In sum, the motion patterns represented by the PC1 and PC2 models are (a) the size of the tongue tip—the largest tip region occurs in the large motion/nontumor slice and the smallest tip region in the small motion/tumor slice; and (b) the direction of motion in the tip, the body, or both. In each slice the PCs reflect these behaviors to different extents, however, so that the "building blocks" of the individual subject's shapes may be different across slices even if the end patterns are similar. These building blocks are artificial constructs of the PCA and indicate slightly more variance in a slice explained by one PC versus another. Although tongue-tip direction and tip size are key components in the apical versus laminal /s/, they did not distinguish the patients from the controls; the use of laminal /s/ was infrequent in both groups.

Classification of Patients Versus Controls Based on PC Data

Two methods were used to determine how well the PCs distinguished between subject groups. First, LDA was performed separately on each slice to determine how many PCs were needed to classify the controls and patients into two distinct groups. In the large motion/nontumor slice and the midline slice, 10 PCs were needed for 100% correct subject classification. Only seven PCs were needed for 100% correct classification in the small motion/tumor slice. Based on Table 2, this means that 92.4% of the variance was enough to distinguish the groups in the small motion/tumor slice, but in the large motion/nontumor and the midline slices, 98.7% and 97.5% were needed, respectively.

Second, *t* tests were used to determine whether any single PCs differentiated between the two subject groups based on significant differences in their loading patterns (see Table 3). Because all PCs are orthogonal, each PC absorbs a percentage of the total variance indicating the size of its effect. In the small motion/tumor side, two PCs approached significance and accounted for 14.7% of the variance: PC3, $p = .093$, 11% variance, and PC7, $p = .055$, 3.7% variance. The effect of both these PCs was to determine whether or not tip and body motion were uniform. PC7 also affected the direction of tip motion. At midline, three PCs approached significance and accounted for 12.8%: PC4, $p = .08$, 6%; PC5, $p = .057$, 5%; and PC10, $p = .09$, 1.5%. In the large motion/nontumor side, PC9 was significantly different between patients and controls, $p = .019$, 1.4%. The effect of ±PC9 was a slight difference in the direction of tongue body velocity, from more forward to more downward.

In sum, several PCs had significantly different loadings for patients versus controls, and revealed key distinctions in their variance. The component features revealed by the PCs were tongue-tip direction and independence in the small motion/tumor slice, tongue-tip and -body direction at midline, and tongue-body direction in the large motion/nontumor slice. In the large motion/nontumor slice, the significant PCs accounted for the least variance (1.4%). In the small motion/tumor slice they accounted for the most variance (14.7%).

Motion Patterns of Controls Versus Patients

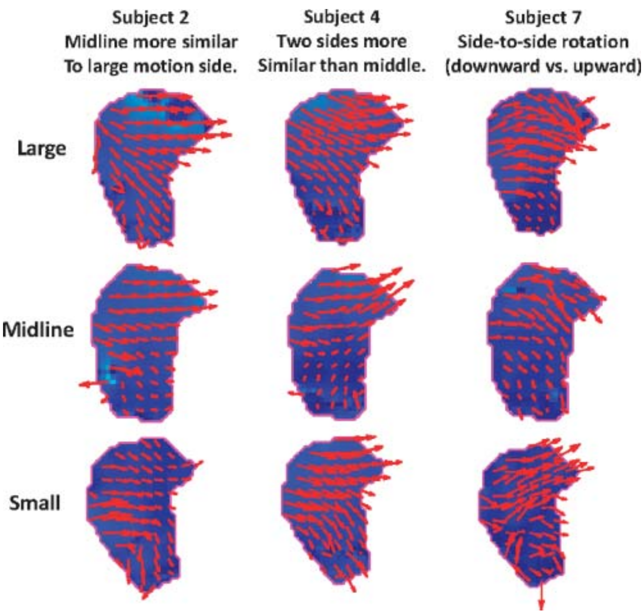
Direct inspection of the velocity fields revealed some similarities between controls and patients. Figure 3 shows data from three control subjects. Subject 2 demonstrated a link between the motion pattern of the midline and one side. The small motion side differed in direction and amount of motion from the other two slices. Subject 4 showed greater similarity between the two sides than the middle, especially in the tip. The forward-moving tip on the sides was in direct contrast to the upward motion at midline. Subject 7 appeared to use left-to-right rotation, because the anterior tongue elevated on the small motion side and lowered in the midline and opposite side. These three subjects reveal a wide range of variability among the controls.

The patient motion patterns can often be categorized grossly in the same terms as those of the controls (see Figure 4). For example, Subject 11 was asymmetrical and had more

Table 3. PCs with significant differences between control and patient loadings.

	PC					
Tissue slice	PC9	PC4	PC5	PC10	PC3	PC7
Large motion/nontumor						
Control (mean loading)	0.72					
Patient (mean loading)	–2.33					
p	.019					
% variance	1.40					
Midline						
Control (mean loading)		0.98	–1.02	–0.49		
Patient (mean loading)		3.30	3.37	1.63		
p		.08	.06	.09		
% variance		5.90	5.40	1.50		
Small motion/tumor						
Control (mean loading)					–1.08	0.69
Patient (mean loading)					3.57	–2.33
p					.093	.055
% variance					11.00	3.70

Figure 3. Individual velocity fields for three control subjects in all three slices.



similar patterns of motion between the nontumor slice and the midline slice than in the tumor slice. Subject 12 had a greater anterior pull on the posterior tongue in the midline than in the two sides. Other features, however, were notably different in the patients. For example, Subject 13 had an in-plane rotation in the upper tongue on the tumor side that propagated laterally, as seen in a lesser rotation at midline and inconsistent velocities in the large motion/nontumor side.

Figure 4 shows plots of the PC1 × PC2 loadings of all subjects in each slice. The extreme loadings of the patients on PC1 and PC2 means that the PCs were influenced by them and the other extreme controls. The PC plots show that the patients (Subjects 11, 12, and 13) often loaded at the extreme limits of PC1 and/or PC2, indicating they represented extreme values of the velocity patterns seen in the other subjects. Subject 11, who had an apical /s/, had a strong negative loading on PC1 in the nontumor and midline slices (see Figure 4). A negative PC1 corresponds to a large tongue tip, indicating that this was the distinguishing feature for this subject, rather than, for example, the more anterior direction of the tongue body also seen in Figure 4. In the midline, his negative loading on PC1 reflected a fairly downward tongue-body motion, whereas the strong positive loading on PC2 reflected the protruded tip. Subject 12, who had a laminal /s/, had notably downward and forward motion in the tongue tip and body for all three slices. The loadings for Subject 12 indicate that in some cases the velocity pattern was similar to the controls, such as in the large motion/nontumor and midline slices, whereas it was more unusual in the small motion/tumor slice. Subject 13, an apical /s/ speaker, had PC plots whose loadings were extreme for PC1 for all three slices, which indicated that the distinguishing component was the lack of tip extension.

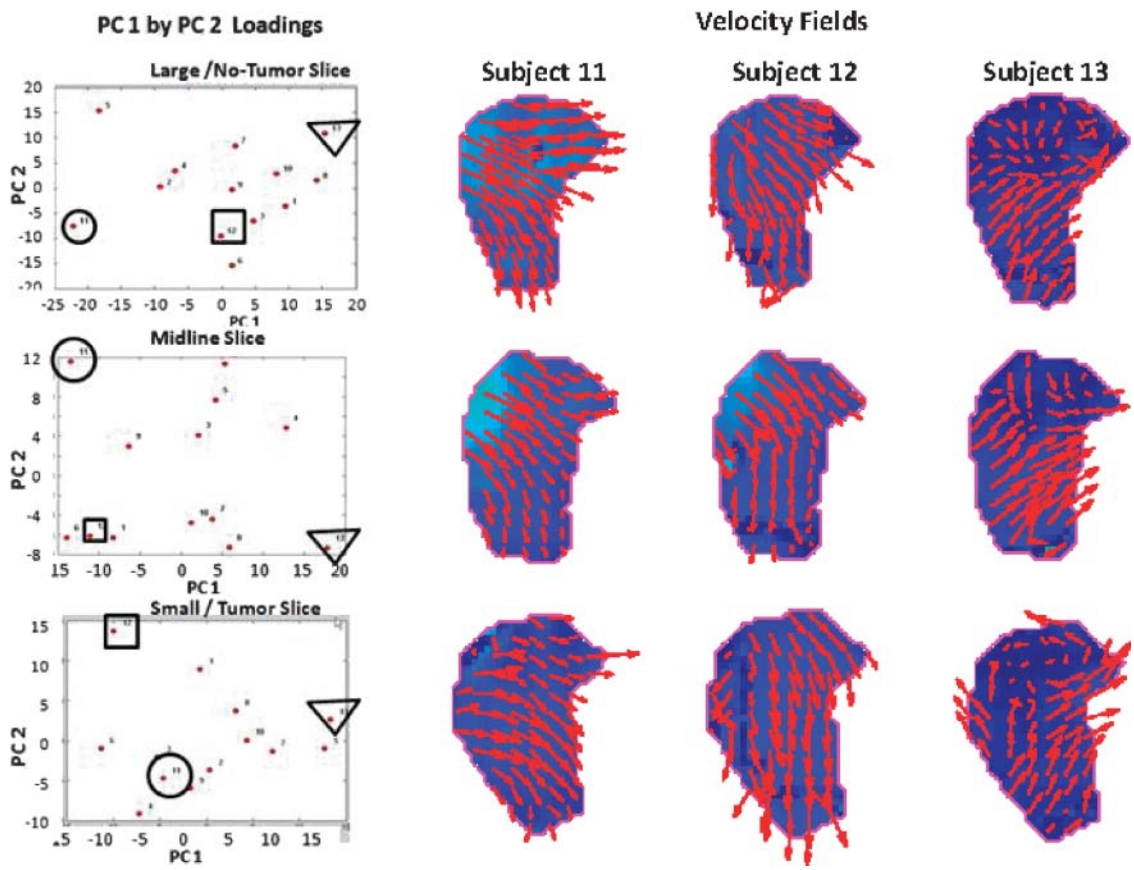
Discussion

The goal of the present study was to quantitatively represent internal tongue motion patterns due to glossectomy surgery, and distinguish them from typical motion patterns. The PCs identified the key components of the tip motion as variance in direction and size of the tip. It was hypothesized that on the tumor side, patient motion patterns would be small in scale and unusual in patterning because of scar tissue restrictions. A second hypothesis was that the nontumor side would use very large motions to reach the palate in compensation for the resected side.

Support for the first hypothesis was found in the small motion/tumor slice. The tongue tip itself is smaller in this slice than in the midsagittal or large motion/nontumor slices, as seen in Figure 2, PC1 models. However, only one patient (Subject 13) loaded very highly on this PC. This slice also showed greater motion pattern differences between patients and controls than the other slices. Only seven PCs were needed to classify 100% of the subjects into their correct groups, whereas 10 PCs were needed to classify subjects in the other slices. The first seven PCs, however, did not capture all the critical features even in the small motion/tumor slice. Subject 13 had an unusual backward rotation that was captured best by PC10. This backward rotation has been observed in other patients and appears to be centered around the scar. It is likely that the scar, which is rigid and less mobile than the rest of the tongue, may not stretch and move sufficiently for typical tongue-tip extension. The backward rotation of Subject 13's tongue elevated the anterior tongue (tip and blade) and in this case produced an apical /s/ contact pattern, an unusual /s/ type for glossectomy patients in general and contrary to our third hypothesis. It is possible that rotation is an adaptive mechanism that allowed elevation of the anterior tongue despite scar rigidity.

The second hypothesis was that the nontumor side would use very large motions to reach the palate as compensation for the limited motion of the resected side. This hypothesis was not supported by the data. On the contrary, in the large motion/nontumor slice only one PC was significantly different between groups, and it accounted for very little variance (1.4%). In addition, 10 PCs were needed to differentiate between the subject groups. Both these results indicate that the patients and controls had very similar velocity patterns in the large motion/nontumor slice, and they were not differentiated well by their patterns of variance. Thus the patients did not use unusual motions in the preserved side to supplement the tumor side. It is possible that they did not need to, because the surgical closures were not accompanied by adhesions and the motility of the tumor side was quite good. Lateral measures might show that sufficient palatal contact for /s/ was achieved by the tumor side, or that contact with the inner aspect of the teeth on the tumor side was sufficient to produce an adequate /s/. The similarity with the controls also means that the nontumor

Figure 4. PC1 × PC2 plots for all subjects in all three slices. Symbols indicate patients: Circle = Subject 11; square = Subject 12; triangle = Subject 13. Right: Individual velocity fields for the patients in all three slices.



side was not strongly restricted by any mechanical limitations of the scar on the tumor side.

The third hypothesis was that the patients would produce a laminal /s/ due to reduced motor control of the tongue tip. In fact, only four subjects used a laminal /s/: Subjects 5, 8, 10, and 12, which includes one patient. The /s/ types were determined by direct inspection of the data sets, using the definitions by Dart (1991). Although /s/ type did not distinguish subject groups, PCs 1 and 2 showed tongue-tip elevation to be a large source of variation between individual subjects. Figure 4 shows Subject 12's tip and body motions to be highly correlated and decidedly downward, with PC loadings that were negative or zero on PCs 1 and 2. Subject 11, who was apical, had forward, but not upward tip motion and was always negative on PC1, suggesting downward motion, but positive on PC2 in the nontumor slice, suggesting upward tip motion. Moreover, this subject may have been apical before the surgery and was able to maintain that pattern due to the nontumor side. A pattern of left–right asymmetry is consistent with some typical controls who also produce unilateral upward motion (see Figure 3, Subject 7) and may be an easy accommodation to make without extraordinary compensation in the nontumor side. Subject 13, however, had the highest positive loadings on PC1 and elevated the tip. Figure 4 indicates that local tip elevation was executed by backward rotation. This unusual behavior suggested that apical /s/ made by patients may be executed in a different manner from controls.

The variability seen among the subjects is due to both anatomical and functional factors. Well-known anatomical contributors to tongue behavior are palate height (Hasegawa-Johnson, Pizza, Alwan, Cha, & Haker, 2003), number and location of teeth (Bankson & Byrne, 1962), tongue size relative to oral cavity (Oliver & Evans, 1986), and dimensions of the vocal tract (anterior–posterior vs. superior–inferior; Mays, Palmer, & Kuhlemeier, 2008). Behavioral contributors would include asymmetrical motion such as left-to-right rotation (Stone, 1990) and asymmetrical tongue contact during /s/. Features that distinguished patients from controls were subtle and tended to occur locally, in the tumor side. The PCs identified the key components of the tip motion as variance in direction and size of the tip. In the patients, in whom innervation of the tip is reduced, tongue-tip size could relate to independence of tip motion from the blade and body. Subject 11 had a "large" tip, suggesting the tip and blade were controlled as a single unit for the /s/ consistent with a laminal gesture, despite the apical appearance. Additional patient data will improve interpretation of these PC results. Full 3-D analysis also will help determine distinctions between groups.

The present study contained several inherent strengths and limitations. The study used PCA to specify the local patterns of motion within the three sagittal tongue slices. The advantage of this method is that it is data-based, not model-based, so the PCs represent components specific to this data set. Thus it was possible to distinguish the similarities and differences among all patients and between the patients and controls. This method also contains several limitations. PCs developed here may not well describe other subjects. Our next plan is to use this method on a large number of control subjects and patients independently and together to more finely detail the groups. Other limitations of this study include the age differences between the patient and the control groups, the small patient-sample size, and the torus and missing teeth for Subject 13, all of which could bias the results. We also cannot determine whether patients changed their /s/-type after surgery because healthy motion MRI scans were not available for the patients. Despite these limitations, the study uncovered new information about the compensatory strategies used post–glossectomy surgery. This set of PCAs was very useful in identifying the two main sagittal dimensions of variance in the motion from /i/ to /s/: the size of the tongue-tip region, and the direction of motion of the tongue tip and body. This was important because the velocity fields themselves contain so much information that it was difficult to determine which features were the most important in distinguishing the groups. For example, the graphs of the PC1 × PC2 loadings were unable to clearly separate the patients from the controls. However, the significant PC loadings in Table 3 clearly separated them. This indicates that subtle differences can be captured by higher order PCs. These results add to our understanding of how small resections allow various adaptive motion strategies to provide excellent recovery of speech.

Acknowledgments

This research was supported in part by National Institutes of Health Grant R01 CA133015, awarded to the first author. Parts of this article were included in the second author's master's thesis. Parts of this article were presented at the 162nd meeting of the Acoustical Society of America, October/November 2011.

References

American Cancer Society. (2014). *Cancer Facts & Figures 2014* [Annual report]. Retrieved from www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2014/index

Axel, L., & Dougherty, L. (1989). MR imaging of motion with spatial modulation of magnetization. *Radiology, 171,* 841–845.

Bankson, N. W., & Byrne, M. C. (1962). The relationship between missing teeth and selected consonant sounds. *Journal of Speech and Hearing Disorders, 27,* 341–348.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer (Version 5.2) [Computer software]. Retrieved from www.praat.org/

Bressmann, T., Ackloo, E., Heng, C.-L., & Irish, J. (2007). Quantitative three-dimensional ultrasound imaging of partially resected tongues. *Otolaryngology—Head and Neck Surgery, 136,* 799–805.

Bushong, S. (2003). *Magnetic resonance imaging: Physical and biological principles* (3rd ed.). St. Louis, MO: Mosby.

Dart, S. N. (1991). Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics, 79,* 1–151.

Dart, S. (1998). Comparing French and English coronal consonant articulation. *Journal of Phonetics, 26,* 71–94.

Fisher, S. E., McKinnon, G. C., Maier, S. E., & Boesiger, P. (1993). Improved myocardial tagging contrast. *Magnetic Resonance in Medicine, 30,* 191–200.

Greene, F. L., Page, D., Morrow, M., Balch, C., Haller, D., Fritz, A., & Fleming, I. (2002). *AJCC cancer staging manual* (6th ed.). New York, NY: Springer.

Hasegawa-Johnson, M., Pizza, S., Alwan, A., Cha, J. S., & Haker, K. (2003). Vowel category dependence of the relationship between palate height, tongue height, and oral area. *Journal of Speech, Language, and Hearing Research, 46,* 738–753.

Logemann, J. A., Pauloski, B. R., Rademaker, A. W., McConnel, F. M., Heiser, M. A., Cardinale, S., … Johnson, J. (1993). Speech and swallow function after tonsil/base of tongue resection with primary closure. *Journal of Speech and Hearing Research, 36,* 918–926.

Masaki, S., Tiede, M., Honda, K., Shimada, Y., Fujimoto, I., Nakamura, Y., & Ninomiya, N. (1999). MRI-based speech production study using a synchronized sampling method. *The Journal of the Acoustical Society of Japan, 20,* 375–379.

Mays, K. A., Palmer, J. B., & Kuhlemeier, K. V. (2008). Influence of craniofacial morphology on hyoid movement: A preliminary correlational study. *Dysphagia, 24,* 71–76.

Michiwaki, Y., Schmelzeisen, R., Hacki, T., & Michi, K. (1993). Functional effects of a free jejunum flap used for recontruction in the oropharyngeal region. *Journal of Craniomaxillofacial Surgery, 21,* 153–156.

NessAiver, M., & Prince, J. L. (2003). Magnitude image CSPAMM reconstruction (MICSR). *Magnetic Resonance in Medicine, 50,* 331–342.

Nicoletti, G., Soutar, D. S., Jackson, M. S., Wrench, A. A., Robertson, G., & Robertson, C. (2004). Objective assessment of speech after surgical treatment for oral cancer: Experience from 196 selected cases. *Plastic and Reconstructive Surgery, 29,* 376–381.

Oliver, R. G., & Evans, S. P. (1986). Tongue size, oral cavity size and speech. *The Angle Orthodontist, 56,* 234–243.

Osman, N. F., Kerwin, W. S., McVeigh, E. R., & Prince, J. L. (1999). Cardiac motion tracking using CINE harmonic phase (HARP) magnetic resonance imaging. *Magnetic Resonance in Medicine, 42,* 1048–1060.

Parthasarathy, V., Prince, J. L., Stone, M., Murano, E. Z., & NessAiver, M. (2007). Measuring tongue motion from tagged cine-MRI using harmonic phase (HARP) processing. *The Journal of the Acoustical Society of America, 121,* 491–504.

Pauloski, B. R., Logemann, J. A., Colangelo, L. A., Rademaker, A. W., & Esclamado, R. (1998). Surgical variables affecting speech in treated patients with oral and oralpharyngeal cancer. *Laryngoscope, 108,* 908–916.

Rastadmehr, O., Bressmann, T., Smyth, R., & Irish, J. (2008). Increased midsagittal tongue velocity as indication of articulartory compensation in patients with lateral partial glossectomies. *Head & Neck, 30,* 718–727.

Stone, M. (1990). A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data. *The Journal of the Acoustical Society of America, 87,* 2207–2217.

Stone, M., Liu, X., Chen, H., & Prince, J. L. (2010). A preliminary application of principal components and cluster analysis to internal tongue deformation patterns. *Computer Methods in Biomechanics and Biomedical Engineering, 13,* 493–503.

Stone, M., & Lundberg, A. (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *The Journal of the Acoustical Society of America, 99,* 3728–3737.

Yorkston, K. M., & Beukelman, D. R. (1981). Communication efficiency of dysarthric speakers as measured by sentence intelligibility and reading rate. *Journal of Speech and Hearing Disorders, 46,* 296–301.

Zhou, X., Stone, M., & Espy-Wilson, C. Y. (2011). A comparative acoustic study on speech of glossectomy patients and normal subjects. In P. Cosi, R. De Mori, G. Di Fabbrizio, & R. Pieraccini (Eds.), *Proceedings of Interspeech 2011: 12th Annual Conference of the International Speech Communication Association, Florence, Italy, August 27–31* (pp. 517–520). Baixas, France: International Speech Communication Association.