

Complexity of vocal tract shaping in glossectomy patients and typical speakers: A principal component analysis

Christina Hagedorn,^{1,a)} Jangwon Kim,² Uttam Sinha,³ Louis Goldstein,⁴ and Shrikanth S. Narayanan⁵

¹Linguistics, College of Staten Island–City University of New York, 2800 Victory Boulevard, Staten Island, New York 10314, USA

²Amazon Care, 410 Terry Avenue North, Seattle, Washington 98109, USA

³Keck School of Medicine, University of Southern California, 1975 Zonal Avenue, Los Angeles, California 90033, USA

⁴Linguistics, University of Southern California, 3601 Watt Way, Los Angeles, California 90089, USA

⁵Viterbi School of Engineering, University of Southern California, 3650 McClintock Avenue, Los Angeles, California 90089, USA

ABSTRACT:

The glossectomy procedure, involving surgical resection of cancerous lingual tissue, has long been observed to affect speech production. This study aims to quantitatively index and compare complexity of vocal tract shaping due to lingual movement in individuals who have undergone glossectomy and typical speakers using real-time magnetic resonance imaging data and Principal Component Analysis. The data reveal that (i) the type of glossectomy undergone largely predicts the patterns in vocal tract shaping observed, (ii) gross forward and backward motion of the tongue body accounts for more change in vocal tract shaping than do subtler movements of the tongue (e.g., tongue tip constrictions) in patient data, and (iii) fewer vocal tract shaping components are required to account for the patients' speech data than typical speech data, suggesting that the patient data at hand exhibit less complex vocal tract shaping in the midsagittal plane than do the data from the typical speakers observed.

© 2021 Acoustical Society of America. <https://doi.org/10.1121/10.0004789>

(Received 5 September 2020; revised 2 April 2021; accepted 6 April 2021; published online 22 June 2021)

[Editor: Susanne Fuchs]

Pages: 4437–4449

I. INTRODUCTION

During speech production, the tongue can be shaped in both gross ways (e.g., tongue body lowering and backing for the production of /a/) and in relatively complex ways (e.g., simultaneously producing anterior and posterior lingual gestures for /ɪ/). Distinct regions of the tongue being differentially controlled to simultaneously form multiple constrictions as is required for the production of certain isolated speech sounds (e.g., /l/ or /ɹ/) gives rise to complex patterns in vocal tract shaping. Further, regardless of the number of simultaneous articulatory gestures required for the production of individual speech segments, natural speech requires ongoing differential control of multiple parts of the vocal tract due to the temporal overlap of articulatory gestures corresponding to neighboring segments (i.e., coarticulation). The loss of lingual tissue and discomfort associated with the glossectomy procedure has been claimed to result in limited lingual mobility (Imai and Michi, 1992; Michi, *et al.*, 1989; Bachher *et al.*, 2002); when the muscular integrity of the tongue is no longer intact, vocal tract shaping due to lingual movement is compromised. Thus, it is expected that post-operatively, patients successfully carry out relatively gross movements of the tongue within the vocal tract to effect changes in vocal tract shaping in broad regions of the vocal tract (e.g., anterior and posterior), but may struggle to move

distinct parts of the tongue simultaneously to effect more complex vocal tract shaping. Many studies have investigated speech articulation following the glossectomy procedure using acoustics (Logemann *et al.*, 1993; Michiwaki *et al.*, 1993; Savariaux *et al.*, 2001; Zhou *et al.*, 2011; McMicken *et al.*, 2012), cineradiography (McMicken *et al.*, 2012), videofluoroscopy (Georgian *et al.*, 1982; Morrish, 1984), electropalatography (EPG) (Fletcher, 1988; Imai and Michi, 1992; Michi, *et al.*, 1989), and cine-MRI (Stone *et al.*, 2004; Stone *et al.*, 2010; Stone *et al.*, 2014a; Stone *et al.*, 2014b). The findings of these and other studies suggest that articulation is least affected in patients who have undergone resection of the base of tongue and that stop consonant articulation is most frequently distorted for patients post-operatively. Past work, also based on acoustic measures, suggests that horizontal tongue movement is generally impacted by the procedure more than vertical tongue movement (Kaipa *et al.*, 2012; Savariaux *et al.*, 2001; McMicken *et al.*, 2012). Some researchers have determined that individuals who have undergone glossectomy and lingual reconstruction exhibit reduced lingual range of motion (Bressmann *et al.*, 2005) and have speculated that lingual flexibility in this patient population is impaired, based on reduced functional independence of distinct tongue segments (Stone *et al.*, 2004). To date, however, no study has set out to compare, through quantitative means, the complexity of vocal tract shaping due to lingual movement in typical and patient populations.

^{a)}Electronic mail: christina.hagedorn@csi.cuny.edu

In this study, we use real-time magnetic resonance imaging (rtMRI) and an analytical method of semi-automatically detecting air-tissue boundaries in vocal tract images to calculate cross-distance measurements along the vocal tract, over time. While several methods (described above) have been used to study post-glossectomy speech, rtMRI is particularly well suited for the investigation of patterns in vocal tract shaping in this patient population. rtMRI provides a dynamic, midsagittal view of the entire vocal tract from the lips to the larynx without exposing the participant to ionizing radiation. Unlike cine-MRI, rtMRI does not require multiple productions for analysis of a single item. In contrast to electromagnetic articulography, rtMRI is minimally invasive in that it does not require adhering sensors to vocal tract articulators that may be tender and still in the process of post-surgical healing at the time of data collection.

In order to index the complexity of vocal tract shaping due to lingual movement in individuals who had undergone glossectomy and in typical speakers, we carry out a Principal Component Analysis (PCA) on vocal tract cross-distance measures. PCA attempts to reduce the number of dimensions needed to represent the data by identifying the most dominant directions of variability in the data. These principal components are ordered such that the first component is most dominant, explaining the most variance in the data, while the last component is least dominant, explaining the least variance in the data. In the current investigation, a PCA is carried out on gridline cross-distance measures, allowing the vocal tract shaping patterns associated with each component to be identified and compared across speakers.

PCA has been used to investigate several biological phenomena, including patterns in brain activation (Suma and Murali, 2007), locomotion (Daffertshofer *et al.*, 2004; Ormoneit *et al.*, 2005; Wang *et al.*, 2013), and cardiac motion (Chandrashekar *et al.*, 2003; Tabassian *et al.*, 2015). PCA has also been applied in speech research, and, in particular, studies of lingual articulation (Bressmann *et al.*, 2005; Bressmann *et al.*, 2007; Stone *et al.*, 2010; Stone *et al.*, 2014a; Stone *et al.*, 2014b). Bressmann *et al.* (2005) and Bressmann *et al.* (2007) used PCA to compare lingual surface shapes, recorded using ultrasound, in patients who had undergone glossectomy with and without reconstruction, pre- and post-operatively, concluding that the surface shape of the tongue is most similar among pre-operative and post-operative patients without reconstruction (two principal components were required to explain their shapes). Explanation of variance in surface shape of the tongue in post-operative patients with reconstructive flaps required an additional principal component, however, reflecting left-right asymmetry and reduced midline concavity. Stone *et al.* (2010) used PCA on tagged cine-MRI data to compare velocity fields in midsagittal tongue slices in typical speakers of multiple languages and an individual who had undergone glossectomy. Later work by Stone *et al.* (2014a) quantified motion differences in individuals who had undergone glossectomy and

typical speakers during a specific speech task involving lingual movement during the transition from /i/ to /s/ using PCA on velocity fields in three sagittal slices of the tongue. This study revealed that the two largest principal components reflected variance in the size and independence of the tongue tip and the direction of motion of the tongue tip, body, or both. Moreover, it was revealed that fewer principal components were necessary to differentiate patient and control data associated with lingual slices in the presence of a tumor than in data associated with non-tumor lingual slices. Further, the degrees of variance reflected by each principal component were found to differ significantly between patients and control speakers. Examination of particular tongue tip patterns exhibited revealed a substantial amount of variability across participants, but suggested that patients' and typical speakers' production of apical /s/ may be achieved using distinct strategies. An extension of this work by Stone *et al.* (2014b) used PCA to determine whether internal tongue motion patterns differ between typical speakers who achieve the production of /s/ in different ways (apically and laminally), and also between typical speakers and individuals who had undergone a glossectomy. This work revealed no discrete difference in internal tongue motion patterns between typical speakers who differ in their production of /s/ (apical vs laminal). Rather, patients who exhibited the apical production patterned uniformly as a group, and markedly differently than all other speakers. Specifically, these speakers exhibited no elevation of the tongue tip, but rather substantial retraction and depression of the tongue tip.

The aim of the current study is to quantitatively identify and compare patterns in vocal tract shaping due to lingual movement during natural speech production in patients who have undergone glossectomy and typical speakers. We predict that for the typical speakers and patients alike, the primary vocal tract shaping pattern, as indicated by the largest principal component, will primarily reflect changes in aperture in the palatal and pharyngeal regions, formed by gross, forward-backward movement of the tongue body. This is consistent with factor analytic studies of tongue motion revealing that the forward-backward movement pattern is prevalent in typical speech, evidenced by high correlation between both distal points on the tongue and between vocal tract aperture measures in the anterior and posterior regions of the vocal tract (Harshman *et al.*, 1977; Maeda, 1990). Moreover, forward-backward movement of the tongue body accounting for the largest amount of variance in typical speech data is expected given that when the tongue is tasked with forming sequential constrictions in distinct regions of the vocal tract, it does so through a pivot motion, such that at the target constriction regions of interest (e.g., the palatal and pharyngeal regions during an /i/ to /a/ transition), vocal tract aperture tends to change drastically, while progressively less change in vocal tract aperture is exhibited as the region intermediate these areas is approximated (such that virtually no change in vocal tract aperture would be exhibited in the uvular region, in the case of an /i/ to /a/ transition) (Iskarous, 2005). The next smallest components may reflect

more subtle changes in vocal tract shaping, such as those brought about by tongue tip constriction and raising of the medial tongue body (e.g., for the production of /u/, /o/, or /l/), and are expected to contribute to far less of the overall vocal tract shape in individuals who had undergone glossectomy than in typical speakers. Accordingly, we expect that fewer principal components will be required to explain 99% of the variance in the patient data than in the typical speakers' data, reflecting reduced complexity of vocal tract shaping for the patients, as compared to typical speakers.

II. METHOD

A. Participants

The participants in this study were six individuals who had undergone glossectomy and two typical speakers (Table I). All participants were native speakers of American English. The mean age of the individuals who had undergone glossectomy was 62 (*SD* = 7.96) and the mean age of the typical speakers was 28.5 (*SD* = 3.53). The typical speakers had no history of speech, language, or hearing difficulties. None of the patients received speech or swallowing therapy between the time of surgical treatment and the time of data collection, which took place at least 4 months post-operatively. All patients underwent (i) resection of the tumor, involving the superficial, outermost layers of the tongue at which the carcinoma originates, and deeper layers to the extent that the tumor had grown into these areas, in addition to an adequate margin of healthy tissue and (ii) reconstruction of the lingual tissue using a radial forearm free flap, the sternocleidomastoid, Alloderm, or a combination thereof. None of the patients had dental involvement. Post-operative chemo-radiation therapy was undergone by patients F1 and M1, and post-operative radiation therapy was undergone by patients M2, M3, M4, and M5. As observable in Fig. 1, the lingual mass of patients is often visibly reduced as compared to that of typical speakers, even following reconstruction, in part due to reconstructive flap shrinkage caused by post-operative radiation therapy. Of importance to investigating the impact of reduced lingual function on vocal tract shaping is that radiation and chemo-radiation therapies are also known to cause fibrosis of soft

tissue which is likely to impede movement of the tongue and the structures surrounding it (Koppetsch and Dahlmeier, 2003; Shin *et al.*, 2012). No preoperative speech data from the patients could be collected, given that scheduling surgery immediately following diagnosis took precedence, and that continuous speaking caused discomfort for the patients due to the presence of the tumor. Further, as discussed in Stone *et al.* (2014a) and Stone *et al.* (2014b), and as demonstrated by Zhou *et al.* (2011), speech samples obtained from patients pre-operatively are likely not representative of the patients' typical speech patterns, due to the structural perturbation of speech by the tumor and the discomfort associated with its presence.

B. Stimuli

The stimuli consisted of select sentences from the TIMIT corpus (Wrench and William, 2000) and the Rainbow Passage (Appendix), made visible to the participants using a projection and mirror setup.

C. Procedure

Image data for all speakers were acquired on a 1.5 T GE Sigma scanner, using a 13-interleaf spiral gradient echo pulse sequence [TR = 6.004 ms, FOV = 200 × 200 mm, flip angle = 15° (20° for M3)] and a custom 4-channel head and neck receiver coil. The scan plane (5 mm slice thickness) was located midsagittally; pixel density in the sagittal plane was 84 × 84 yielding a resolution of 2.38 × 2.38 mm. Image data were acquired at a rate of 12.8 frames per second, and reconstructed at 23.79 frames per second using a sliding window technique. Audio was recorded inside the scanner at 20 kHz simultaneously with the MRI acquisition, and subsequently noise reduced (Bresch *et al.*, 2006).

D. Articulator segmentation of real-time MRI data and vocal tract cross-distance measurement

First, intensity-corrected frames (Kim *et al.*, 2014) were created by multiplying the pixel intensities of the original frames by the inverse of the pixel intensities smoothed by two-dimensional median filtering (5 × 5 pixels). The final images were formed by using sigmoid-kernel-based

TABLE I. Glossectomy patient and typical speaker data.

		Age	Locus of resection	Tumor size	Reconstruction	Post-operative therapy	Duration post-op
PATIENTS	F1	52	(Bilateral) Base of tongue extending into oral tongue	T3 (4–6 cm)	Yes—radial forearm free flap	Chemo-radiation therapy	> 6 mos.
	M1	70	Oral tongue	T4 (>6 cm)	Yes—radial forearm free flap	Chemo-radiation therapy	6 mos.
	M2	67	Base of tongue	T3 (4–6 cm)	Yes—radial forearm free flap	Radiation therapy	4 mos.
	M3	55	Base of tongue	T3 (4–6 cm)	Yes—sternocleidomastoid muscle and Alloderm	Radiation therapy	> 6 mos.
	M4	66	Base of tongue	T3 (4–6 cm)	Yes—Alloderm	Radiation therapy	> 6 mos.
	M5	55	(Unilateral: left) Oral and base of tongue	T2 (2–4 cm)	Yes—sternocleidomastoid muscle and Alloderm	Radiation therapy	> 6 mos.
TYPICAL SPEAKERS	FT1	26	—	—	—	—	—
	MT1	31	—	—	—	—	—

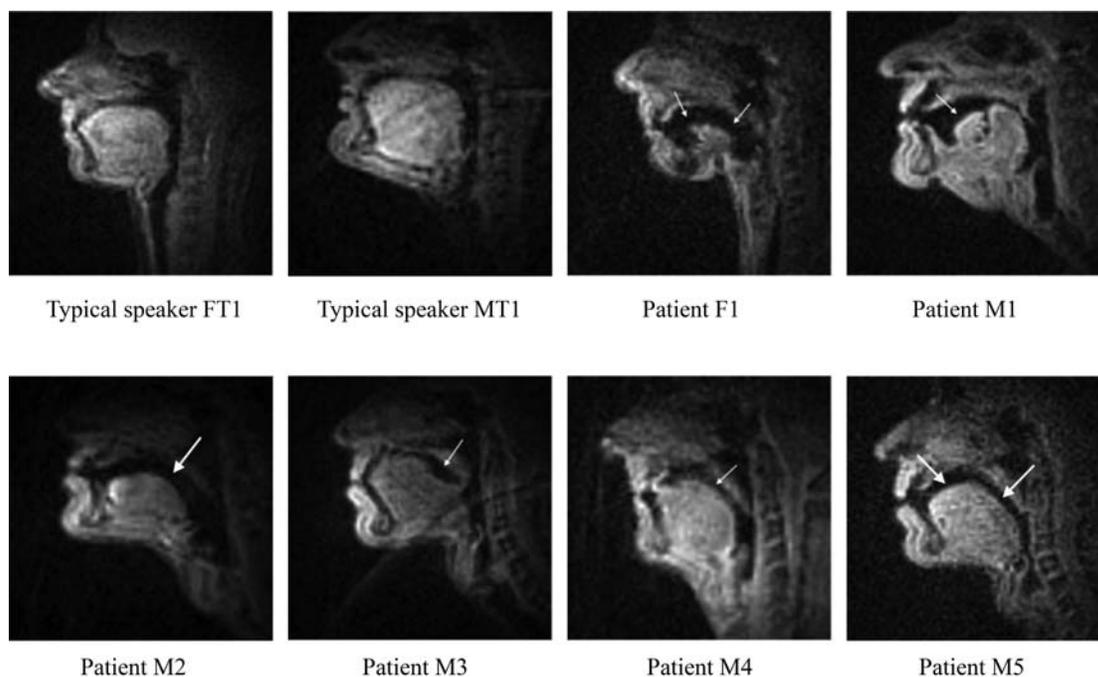


FIG. 1. Real-time MRI images of participants at rest. Top: Typical speakers FT1 and MT1, patient F1 (base of tongue extending into oral tongue), patient M1 (oral tongue). Bottom: patient M2 (base of tongue), patient M3 (base of tongue), patient M4 (base of tongue), patient M5 (unilateral oral and base of tongue). Arrows indicate loci of resection and reconstruction.

intensity warping within each frame. This process increased the contrast between high intensity pixels and low intensity pixels even further. The midline of the vocal tract was then identified semi-automatically [details in Kim *et al.* (2014)], based on the manual selection of speaker-specific anatomical landmarks including the inter-labial, palatal, and laryngeal regions (Fig. 2).

Then, all frames except those associated with acoustic silence were fit with gridlines (Fig. 3) along the vocal tract from the lips to the larynx. This set of static gridlines was placed on the vocal tract midlines such that the slope of each gridline was perpendicular to the midline. Distances of adjacent gridline points were set to be equal on the midline.

Air-tissue boundaries were identified by locating and connecting intensity thresholds along the vocal tract gridlines [details in Proctor *et al.* (2010)]. For all frames on which articulator segmentation was applied, vocal tract cross-distance measures (in pixels) were recorded for each gridline.

E. Principal Component Analysis

Given that the present analysis intends to address how glossectomy affects the complexity of vocal tract shaping due to loss of typical lingual function, gridlines anterior of the alveolar region were excluded from the PCA. To allow for PCA result comparison across participants, interpolation

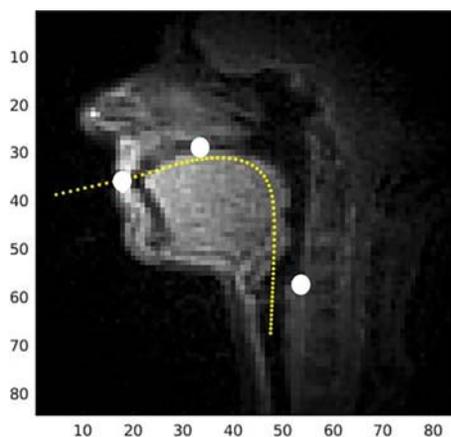


FIG. 2. (Color online) White circles: Manually selected, speaker-specific, labial, palatal, and laryngeal landmarks used to identify vocal tract midline. Yellow: Vocal tract midline used to determine orientation of gridlines.

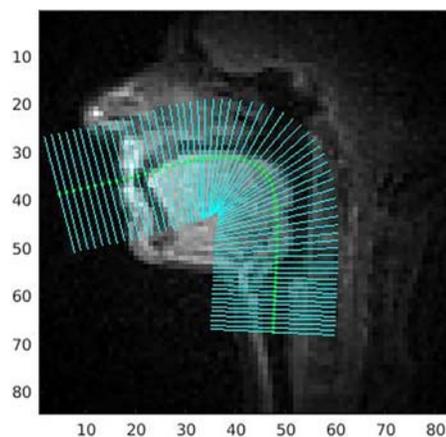


FIG. 3. (Color online) Vocal tract cross-distance gridlines, within which air-tissue boundaries are detected based on intensity thresholds, and vocal tract cross-distance measures are calculated.

TABLE II. Number of components required to explain 99% of the variance in the data; Fewer components are required to explain 99% of the variance in the data for patients (except M5) than typical speakers.

		Area of Resection	Number of components required to explain 99% of variance in data
PATIENTS	F1	Base of tongue extending partially into oral tongue (bilateral)	10
	M1	Oral tongue	12
	M2	Base of tongue	13
	M3	Base of tongue	12
	M4	Base of tongue	12
	M5	Oral and base of tongue (unilateral)	14
TYPICAL SPEAKERS	FT1	—	14
	MT1	—	14

of gridline values was carried out on each participant’s data such that 100 gridlines were associated with the vocal tracts of all speakers, from the alveolar ridge to the larynx. The Principal Component Analysis function (pca) in MATLAB was used to perform principal component analyses on the gridline cross-distance matrices for all speakers, individually. The data matrices analyzed were of size $n \times 100$, where n equals the number of speech frames produced by a given speaker. Columns correspond to the 100 vocal tract gridlines. Gridline cross-distance measure (\hat{X}) can be reconstructed by multiplying the principal component score matrix (Z) by the transposed principal component coefficient matrix (VT), and adding the estimated means of the 100 variables (gridlines) in the original data matrix (μ),

$$\hat{X} = Z * VT + \mu.$$

First, the number of components required to explain 99% of variance in the data for each participant, which was used as an index of the complexity of vocal tract shaping, was compared across speakers. Second, the percentage of variance explained by each of the first three principal components, reflecting the relative contribution of each component to overall vocal tract shaping, was compared across speakers.

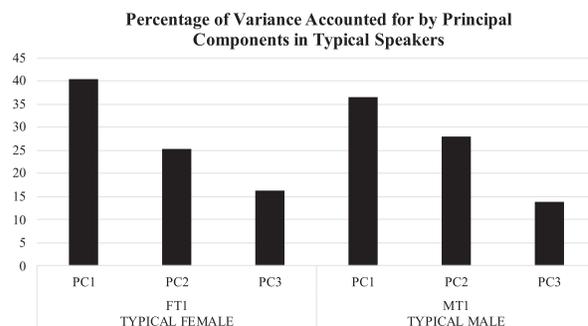


FIG. 4. Percentage of variance accounted for by principal components in typical speakers.

Finally, to illustrate the degree to which each of the first three principal components determines vocal tract shaping at each point along the vocal tract, plots representing the unit positive and negative weightings of the first three factors on the 100 vocal tract gridlines were generated and compared across speakers.

III. RESULTS

As illustrated in Table II, for all patients other than M5, fewer principal components, reflecting distinct movement patterns, are required to capture 99% of the variance in the data, suggesting that the patterns of vocal tract shaping in these patients are less complex than for typical speakers.

As shown in Fig. 4, for typical speakers, the first principal component accounts for ~40% of the overall change in vocal tract shape during speech, while the second and third components account for ~27% and ~14%, respectively.

As illustrated in Fig. 5, for the patient data, with few exceptions, the first component explains a larger amount of variance than in the case of the typical speakers, while the second and third components explain less variance than those of the typical speakers.

For patients M1 and F1, who underwent glossectomy for bilateral oral and base of tongue cancer, and oral tongue cancer, respectively, the first principal component accounts for roughly 60% of the change in vocal tract shape, while the second and third components account for roughly 15% and 7%, respectively (Fig. 5).

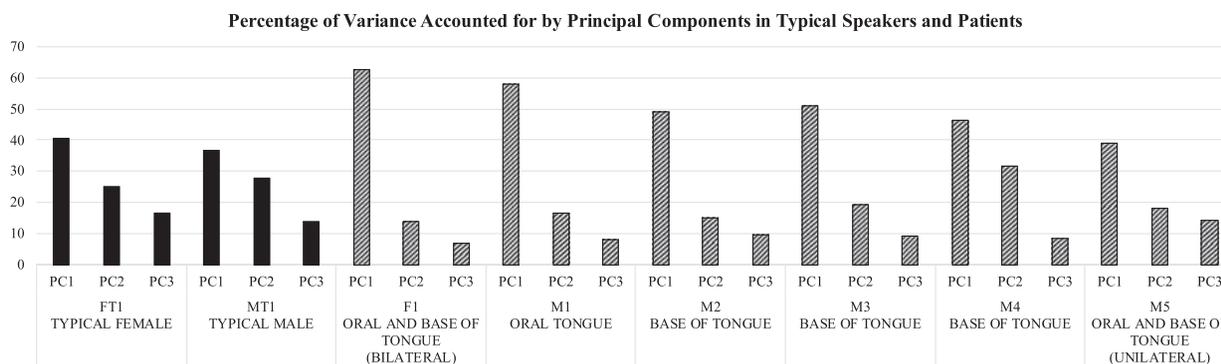


FIG. 5. Percent of variance accounted for by first three principal components in all participants. Overall, patients exhibit higher percentages of variance accounted for by first component and lower percentages of variance accounted for by second and third components than do typical speakers.

In the data from patients who underwent glossectomy for base of tongue cancer, the first component explains roughly 48% of the variance in overall tongue shape; more than accounted for by the first principal component in typical speakers (Fig. 5). For patient M4, the second component values resemble typical values, while for the other patients who underwent glossectomy for base of tongue cancer, the second component explains less variance than in the case of typical speakers. The third component accounts for roughly 10% of the variance; less than in the case of typical speakers.

For patient M5, who underwent glossectomy for unilateral oral and base of tongue cancer, the first component explains 46% of the variance in the data, while the third component explains 14%, patterning similarly to typical speakers. The second component explains 18% of the variance in the data, less than the second component for typical speakers.

The plots displayed in Fig. 6 represent the effect of unit positive and negative weightings of the first three factors on the 100 vocal tract gridlines. The horizontal axis represents gridlines with the alveolar ridge at left and the larynx at right. The vertical axis represents the loading coefficients for the individual components, scaled by their respective eigenvalues (the displacement along that gridline from its mean, with positive values indicating displacement that creates a more constricted shape). Loading coefficients, themselves, index the degree of correlation between the component at hand and specific locations (gridlines) along the vocal tract. While loading coefficients centered on zero indicate that the corresponding variable (gridline) has no bearing on the component at hand, large positive or negative loading coefficients indicate that the corresponding variable strongly influences that principal component. The signs of loading coefficients for a given component are arbitrary; inclusion of opposite-signed coefficients (dashed lines) in the plots allows for visualization of the full range of vocal tract shaping patterns reflected by the principal components. For both the typical male (MT1) and female (FT1) speakers, the loading coefficients of the first principal component (PC1), scaled by their respective eigenvalues, are highest in the palatal and pharyngeal regions of the vocal tract (gridlines ~25 and ~70, respectively) [Figs. 6(a) and 6(d)]. Those corresponding to the second largest principal components (PC2) have the highest amplitude in the velar region (gridline ~50) [Figs. 6(b) and 6(e)]. The amplitude of the third principal components' (PC3) loading coefficients is highest in the region of the alveolar ridge (gridline ~3) [Figs. 6(c) and 6(f)]. The maximum amplitudes of loading coefficient plots across the first three principal components are comparable.

A. Patients F1 and M1: Oral tongue cancer

For patient F1, who underwent glossectomy for oral and base of tongue cancer, the amplitude of scaled loading coefficients for the first principal component is largest in the

palatal and pharyngeal regions (gridlines 17 and 86, respectively) [Fig. 6(g)]. The amplitude of the scaled loading coefficients on the second principal component is uniform amongst most points along the vocal tract, and the maximum amplitude of the scaled loadings on the second component is notably less than that observed for the first component [Fig. 6(h)]. The scaled loading coefficients for the third principal component are distributed rather evenly along the vocal tract, with small peaks in the alveolar and lower pharyngeal regions. The maximum amplitude is lowest among the three component loading plots [Fig. 6(i)].

For patient M1, who underwent glossectomy for oral tongue cancer, the amplitude of scaled loading coefficients for the first principal component is largest in the palatal and pharyngeal regions (gridlines 25 and 75, respectively), though the amplitude in the palatal region is substantially lower than in the pharyngeal region [Fig. 6(j)]. The overall amplitude of the scaled loading coefficients on the second principal component is substantially lower than for the first component, though has maxima in the anterior palatal and lower pharyngeal and laryngeal regions (gridlines 10 and 90, respectively) [Fig. 6(k)]. The overall amplitude of the scaled loading coefficients on the third component is lowest, and values remain uniform along the length of the vocal tract [Fig. 6(l)].

Generally, the scaled loading patterns of patients F1 and M1 differ from those of typical speakers in that the scaled loadings of the second and third principal components are substantially smaller in amplitude than those of the first component, and reflect less change in vocal tract aperture in the velar and alveolar regions.

B. Patients M2, M3, and M4: Base of tongue cancer

The scaled loading patterns exhibited by the patients who underwent glossectomy for base of tongue cancer are distinct from those of both the patients who underwent glossectomy for oral tongue cancer and the typical speakers, yet similar within-group. For patient M2, the amplitude of scaled loading coefficients for the first principal component is largest in the palatal and pharyngeal regions (gridlines 12 and 70, respectively) [Fig. 6(m)]. The overall amplitude of the scaled loading coefficients for the second component is substantially lower than that of the first, and reaches a maximum in the velar region and exhibits a small peak in the alveopalatal region (gridlines 45 and 8, respectively) [Fig. 6(n)]. The amplitude of the coefficients on the third component is lowest, and the plot exhibits local maxima in the alveopalatal and pharyngeal regions (gridlines 8 and 80) [Fig. 6(o)].

For patient M3, the amplitude of scaled loading coefficients for the first principal component is largest in the palatal, lower pharyngeal and laryngeal regions (gridlines 25, 87, and 100 respectively), though highest in the palatal region [Fig. 6(p)]. The overall amplitude of the scaled loadings on the second component is less than those of the first, with maxima exhibited in the alveopalatal and velar regions

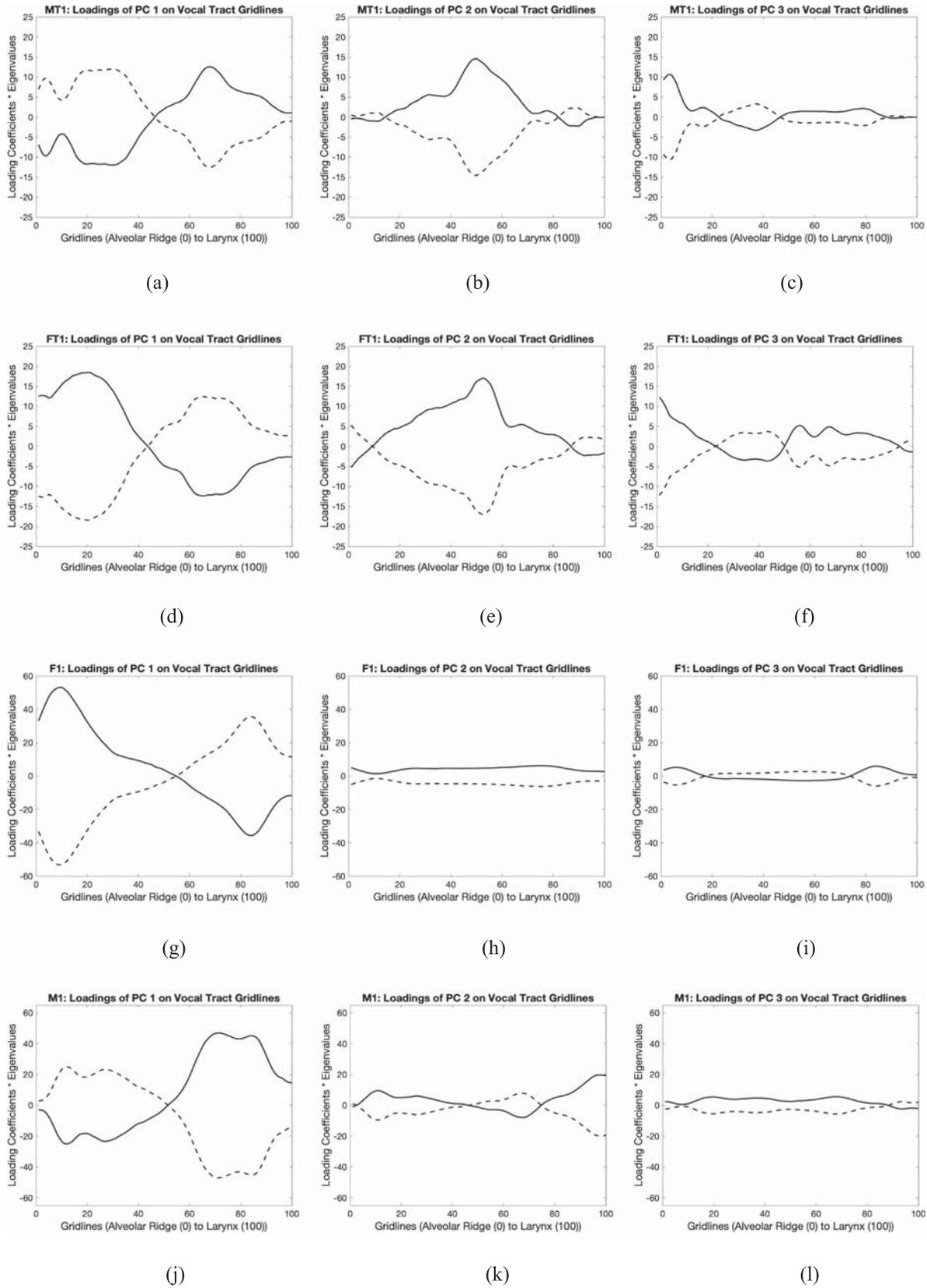


FIG. 6. (Color online) Scaled loadings of principal components 1, 2, and 3 for all speakers. Across speakers, PC 1 values reflect movement in the palatal and pharyngeal regions. PC 2 reflects movement in the velar region for typical speakers FT1 and MT1, and patient M5, movement in both the alveolar/alveopalatal and velar regions for patients M2, M3, and M4, movement in anterior palatal and lower pharyngeal and laryngeal regions with overall low amplitude for patient M1 and does not reflect localized movement at particular regions within the vocal tract for patient F1. PC3 reflects movement in the alveolar/alveopalatal region for typical speakers MT1 and FT1, and patient M5, movement in the alveolar/alveopalatal and pharyngeal regions for patients F1, M2, M3, and M4, and do not reflect localized movement at particular regions within the vocal tract for patient M1.

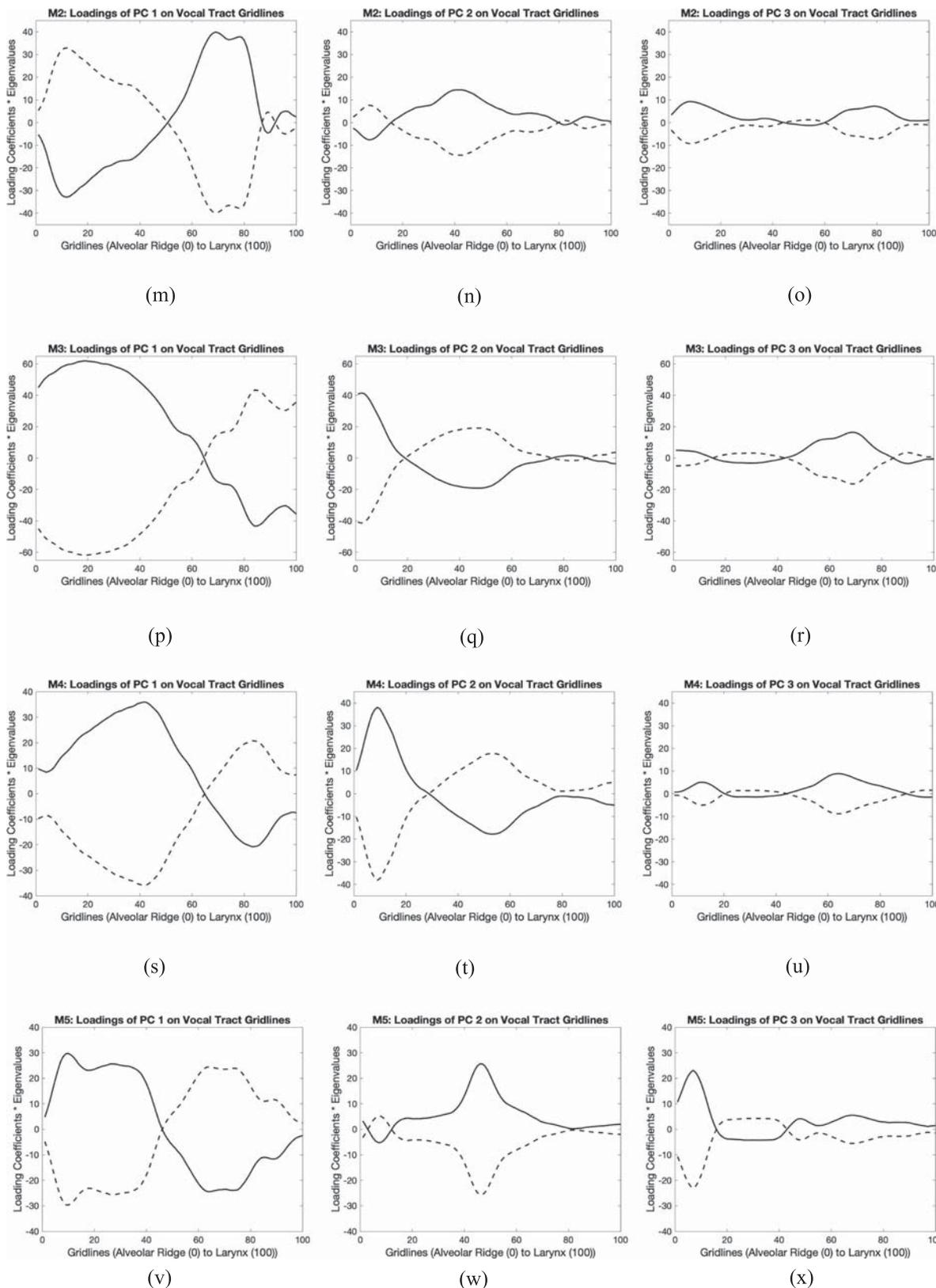


FIG. 6. (Continued).

(gridlines 10 and 45, respectively) [Fig. 6(q)]. The overall amplitude of the scaled loadings on the third component is lowest, with maxima in the alveolar and pharyngeal regions (gridlines 2 and 70) [Fig. 6(r)].

For M4, the amplitude of the scaled loading coefficients for the first principal component is largest in the alveopalatal and pharyngeal regions (gridlines 40 and 84, respectively), though slightly higher in the palatal region

[Fig. 6(s)]. The overall amplitude of the scaled loadings on the second component is slightly lower than for those on the first, and maxima are observed in the alveolar and velar regions (gridlines 9 and 40, respectively) [Fig. 6(t)]. The scaled loadings on the third component are lowest among the three components, and have maxima in the alveopalatal and upper pharyngeal regions (gridlines 10 and 66, respectively) [Fig. 6(u)].

The scaled loadings of patient M5, who underwent unilateral oral and base of tongue cancer, patterned similarly to those of the typical speakers. For patient M5, the amplitude of scaled loading coefficients for the first principal component is highest in the palatal and pharyngeal regions (gridlines 10 and 65, respectively) [Fig. 6(v)]. The scaled loadings on the second component are highest in the velar region (gridline 45) [Fig. 6(w)]. The loadings on the third component exhibit a maximum in the alveopalatal region (gridlines 0–5) [Fig. 6(x)]. Maximum amplitudes of loading coefficients on all components are comparable.

IV. DISCUSSION

The principal components analyses reveal that the number of principal components required to explain 99% of the variance in the data across speakers varies systematically among the patients and typical speakers in this study. The data from typical speakers FT1 and MT1 require more explanatory components than do data from all patients except for patient M5, who underwent glossectomy for unilateral oral and base of tongue cancer. The typical speakers and patient M5 each require fourteen components to explain 99% of the variance in the data, while the patients who underwent glossectomy for base of tongue cancer, and patient M1, who underwent glossectomy for oral tongue cancer require slightly fewer components (12–13). The data from patient F1, who underwent resection involving bilateral oral and base of tongue regions, requires the fewest (10) explanatory components. These findings are largely consistent with the hypothesis that the patient data would require fewer principal components, reflecting less complex vocal tract shaping due to tongue movement, to capture 99% of the variance in the data than the data from typical speakers. That data from patient M5, who underwent glossectomy for unilateral oral and base of tongue cancer, patterns as data from typical speakers is likely due to limited impairment of left-sided lingual mobility, given that the reconstructed left lingual tissue is transported vertically and horizontally (in the midsagittal plane) by the intact right-sided lingual tissue and his tumor being smaller in size (T2) than those of the other patients.

Analysis of the particular loading plots, representing the unit positive and negative weightings of the first three factors on the 100 vocal tract gridlines, reveal that the typical speakers in this study pattern remarkably similarly in the behavior of the three components accounting for the largest amounts of change in vocal tract shape. That the first component accounts for change in the anterior-posterior

dimension of the vocal tract and that the second component accounts for change in the velar region is consistent with previous principal component analyses of lingual movement (Harshman *et al.*, 1977; Maeda, 1990; Iskarous, 2005). The third largest component accounts for vocal tract shape change in the alveolar region, within which tongue tip constrictions are frequently made during running speech in English. Importantly, the scaled loading coefficient plots exhibit similar maximum amplitudes, suggesting that movements of the tongue body into smaller, more specific regions, as evidenced by the scaled loading plots of the second and third component coefficients, contribute just as much to the overall vocal tract shaping as do the gross, forward-backward lingual movements that likely effect the changes in vocal tract shaping reflected by the first component. That the finely controlled lingual movements in these more specific regions of the vocal tract contribute to vocal tract shaping as heavily as do the gross, forward-backward movements of the tongue body in the general anterior and posterior regions is consistent with the hypothesis that typical speakers exhibit a considerable degree of complexity in vocal tract shaping due to lingual movement.

Similar to the case of typical speakers, patient F1's first principal component is associated with changes in vocal tract shaping in the anterior-posterior dimension, reflecting forward-backward movement of the tongue. The second principal component, associated with changes in vocal tract aperture in the velar region in typical speakers, is largely single-signed and is uniform in amplitude along most of the vocal tract. This is not surprising, given that F1 exhibits difficulty forming target velar constrictions, frequently producing compensatory labial constrictions in their place. It is possible that while localized movement at specific gridlines is not reflected by this component, movement that is anatomically localized to the floor-of-the-mouth compensatorily elevating the entire glossectomized tongue, as observed by Xing *et al.* (2019), and consequently affecting aperture at most gridlines along the vocal tract rather uniformly, is reflected. The third principal component, associated with changes in vocal tract aperture in the alveolar region in typical speakers, reflects a particularly small degree of change in aperture in the alveolar and pharyngeal regions for patient F1. That the contribution of the third principal component to vocal tract movement in the alveolar region is substantially lower for patient F1 than for typical speakers is consistent with the locus of her tissue resection. A substantial portion of her oral tongue, comprising the superior and inferior longitudinal muscles responsible for tongue tip movement, has been resected. Interestingly, the principal component score plots associated with the second and third components for patient F1 exhibit much smaller maximum amplitudes as compared to the first component plot and as compared to typical speakers. Amplitude of the loading plots is driven in large part by the eigenvalues associated with each component, indexing the amount of variance accounted for by each component. Thus, the patterns observed suggest that the second and third components, reflecting changes in vocal

tract aperture likely resulting from subtle movements of the tongue, do not contribute as much to the overall shaping of the vocal tract during speech as does the first component, reflecting changes in generally anterior and posterior regions, likely given rise to by gross, forward-backward lingual movement. This is consistent with the hypothesis that some patients exhibit less complex vocal tract shaping, maintaining regular changes in aperture in the anterior and posterior regions of the vocal tract by means of forward-backward lingual movement, but struggling to produce changes in more specific regions of the vocal tract through forming subtler lingual movements.

For patient M1, who underwent glossectomy for oral tongue cancer, the first principal component accounts for changes likely due to forward-backward lingual movement, as was the case for the typical speakers and patient F1. The amplitude of the scaled loading coefficients is markedly lower in the palatal region than in the pharyngeal region, reflecting the presence of a reconstructed flap in the coronal region, and its inability to be finely controlled, which effects less dynamic vocal tract aperture in this anterior region. Accordingly, the loading plot corresponding to the third principal component, reflecting aperture changes in the alveolar regions for typical speakers, does not exhibit a peak in this region, but rather is uniform along the length of the vocal tract. As in the case of the second component for patient F1, it is possible that while localized movement at specific gridlines is not reflected by this component, movement that is anatomically localized to the floor-of-the-mouth compensatorily elevates the entire glossectomized tongue and consequently affects aperture at most gridlines along the vocal tract rather uniformly. The low amplitude coefficient functions associated with the second and third principal components are consistent with the hypothesis that post-operative patients are able to produce forward-backward movement of the tongue body, but tend not to produce subtle lingual movements contributing to dynamic changes in aperture in more specific regions of the vocal tract.

Patients M2, M3, and M4, who underwent glossectomy for base of tongue cancer, exhibit similar patterns in the first component's scaled coefficient functions; aperture changes in the anterior-posterior dimension are captured, with highest amplitudes falling in the palatal and pharyngeal regions. For patients M3 and M4, the maximum amplitude in the palatal region is larger than that in the pharyngeal region, reflecting the specific loci of lingual resection for these patients. Interestingly, for patient M2, the maximum amplitude in the palatal region is lower than in the pharyngeal regions, perhaps contrary to what might be expected given the locus of M2's resection at the base of tongue. It is notable, however, that the length of the posterior vocal tract within which the first component reflects change is short as compared to the length of the palatal region within which change is reflected by this component. Further, it is critical to consider the effects of the glossectomy procedure not only in terms of loss of structure (i.e., the reduction of lingual mass in a particular region of the vocal tract), but also in terms of consequential loss of

function in regions distal to the region in which structure has been lost. It is possible that resection of lingual tissue in the posterior tongue, likely affecting the posterior genioglossus, impedes lingual protrusion required for production of palatal and alveolar constrictions. Additionally, it is conceivable that as a result of posterior lingual tissue resection, patients may compensatorily shift the entire residual tongue posteriorly during speech in order to facilitate production of posterior constrictions, resulting in an overall reduction of lingual tissue occupying the anterior oral cavity. All patients who underwent glossectomy for base of tongue cancer pattern similarly in that the second principal component reflects movement in the alveolar/alveopalatal and velar regions, in contrast to typical speakers, for whom only vocal tract aperture changes due to velar movement was reflected. For all patients who underwent glossectomy for base of tongue cancer, the third principal component captures aperture change in the alveolar/alveopalatal and pharyngeal regions, whereas the third component for typical speakers reflects aperture change strictly in the alveolar region. Consistent with patterns exhibited by M1 and F1, the second and third components' loading scores for the patients who underwent glossectomy for base of tongue cancer are lower in amplitude than those for the first component and those produced by typical speakers. This suggests that for the patients who underwent glossectomy for base of tongue cancer, as well, the components reflecting aperture changes in very specific regions of the vocal tract, likely given rise to by subtle lingual movements, contribute less to the overall vocal tract shaping than for typical speakers, and is consistent with the fact that the percentage variance explained by the second and third components is relatively less than in typical speakers.

The loading plots of patient M5, who underwent glossectomy for unilateral oral and base of tongue cancer, exhibit patterns strikingly similar to those of typical speakers, with the first component reflecting vocal tract aperture changes in the palatal and pharyngeal regions, the second component reflecting changes in the velar region, and the third component reflecting changes in alveolar region. These observed similarities in vocal tract cross-distance patterns are consistent with data from M5 patterning similarly to that of typical speakers in terms of number of principal components required to explain 99% variance in the data. As suggested above, this likely arises due to the unimpaired lingual tissue on the right side of the tongue effectively transporting the reconstructed tissue on the left side of the tongue vertically and horizontally in the midsagittal plane.

Taken together, the vocal tract shaping patterns represented by each principal component (Fig. 6), and the degree to which each of the first three principal components contributes to overall vocal tract shaping (Figs. 4 and 5) indicate that the typical speakers, the patient who underwent bilateral resection of the oral and base of tongue, the patients who underwent bilateral resection of the base of tongue, the patient who underwent bilateral resection of the oral tongue, and the patient who underwent unilateral resection pattern remarkably alike within-groups, and quite differently from each other.

Generally, for the typical speakers in this study, the gross, forward-backward movement pattern associated with the first component contributes to the overall tongue shaping pattern proportionally less than it does for individuals who had undergone glossectomy. For typical speakers, the subtler vocal tract shaping patterns associated with the second and third components contribute proportionally more to the overall shaping of the vocal tract than for the patients who underwent bilateral glossectomy. The first and second principal components associated with patient M4's data accounts for only slightly more variance in the data than those of typical speakers, however, while the third component accounts for very little variance (consistent with the results for other patients). Consistent with the patterns exhibited by the loading plots, the first and third principal components for data from patient M5, who underwent glossectomy for unilateral oral and base of tongue cancer contribute to overall vocal tract shaping to a similar degree as do those of typical speakers. The second principal component, reflecting aperture changes in the velar region, however, contributes less to overall shaping than that of typical speakers.

In sum, the data at hand align with the hypothesis that patients who have undergone glossectomy exhibit less complex vocal tract shaping than do typical speakers, in the following ways.

- (i) Fewer principal components, reflecting patterns of vocal tract constriction, are required to account for data from the patients that underwent glossectomy for oral and base of tongue cancer in our study than for data from the typical speakers.
- (ii) The first principal component, reflecting aperture changes in broadly anterior and posterior regions of the vocal tract, accounts for a disproportionately large amount of the overall movement in the vocal tracts of the patients as compared to the typical speakers.
- (iii) Other principal components, reflecting changes in vocal tract shape due to subtle lingual movements in more specific regions of the vocal tract (i.e., alveolar and velar regions), account for far less variance in the patient data than in the data from typical speakers, and pattern expectedly given the particular loci of the patients' resections.

The results of this study are comparable to those of [Bressmann *et al.* \(2007\)](#) only to a limited extent, given differences in methodologies, and therefore what the principal components represent. However, [Bressmann *et al.* \(2007\)](#) found that post-operatively, the amount of variance in the data explained by the first principal component increased in the case of patients with no reconstruction (as we found for patients in our study, all of whom underwent reconstruction), and remained largely the same in the case of patients who received reconstructive flaps. Likewise, for both groups of patients, the amount of variance accounted for by the second principal component decreased post-operatively, with which the findings of the present study are consistent.

Though [Bressmann *et al.* \(2007\)](#) reports on three explanatory patterns having emerged in the case of the post-operative data from the patient with flap reconstruction, the number of principal components required to explain a very high percentage of the data (e.g., 99%) for each patient is not reported. Thus, comparison of the findings of the two studies in terms of the number of components required to explain the patient and typical data is not possible.

While the study at hand provides new insight into how patterns of vocal tract shaping in typical speakers and patients who have undergone glossectomy may differ, is not without limitations, and thus paves the way for future directions of this work. Though the relatively small sample populations used in our study enable investigation and comparison of subtleties of within-speaker patterns, they limit the statistical power of the study, and thus limit the generalizability of our findings to the patient population at large. Though the PC patterns revealed by our study are consistent with our hypothesis, it is possible that our observation of fewer components being required to explain the patient data than typical data emerges due to random fluctuations in the number of required principal components expected to occur across speakers. Additionally, it is unclear whether the degree of variability observable in the typical data examined in this study, based on only two participants, is representative of the degree of variability that would be observed in a larger sample of typical speakers. Inclusion of a larger sample of patients and typical speakers will allow for (i) a more robust comparison of vocal tract shaping and the number of principal components, reflecting movement patterns, required to explain the data and (ii) generalizations to be made regarding lingual movement tendencies that may provide insight into spontaneously developed compensatory mechanisms (i.e., those not adopted under the direction of a clinician) of patients who have undergone different types of resection. Continuing research may also utilize the analytical methods used in this study to compare patterns in vocal tract shaping within individuals preceding and following surgical tumor resection. Comparison to typical speakers may, however, still be necessary, given that the pre-operative speech of patients may not reflect patterns produced by typical speakers, as a result of the additional mass and physical discomfort caused by the tumor ([Zhou *et al.*, 2011](#); [Stone *et al.*, 2014a](#); [Stone *et al.*, 2014b](#)).

Additionally, the average age of the patient participants in this study was higher than that of the control group, given that the inclusion criteria for participation in MRI studies tend to select for relatively young experimental participants who do not have metal inside the body (e.g., in the form of surgical plates or rods, mesh implants, or cardiac pacemakers). Though the existing literature suggests that labial gestures are most likely to undergo substantial articulatory changes over the course of the lifespan, rather than lingual gestures affecting vocal tract shaping posterior to the teeth (on which this study focuses) ([Bilodeau-Mercure and Tremblay, 2016](#)), future studies would benefit from the inclusion of an age-matched control group to rule out this potentially confounding variable.

Further, in attempt to collect phonetically balanced speech samples that were maximally natural in production, participants in this study were asked to read the stimuli at a comfortable speaking rate. This resulted in a difference in average speech rate between the patient group (3.1 syllables per second) and control group (4.25 syllables per second) which may have contributed to the observed differences in PCA patterns for each group. As observed in Table III, while the speech rates of patients M4 and M5 are identical, the patterns exhibited by each are quite distinct in ways that likely reflect other differences between the two patients, including locus of resection and tumor size, suggesting that although it is possible that speaking rate differences contributed to the PCA differences observed, it is unlikely to have served as a primary modulator of the exhibited patterns. Future investigation of these patterns would benefit from inclusion of both highly natural, self-paced speech samples, as well as samples in which speech rate is controlled.

Additionally, although the current study focuses on the complexity of vocal tract shaping due to lingual movement, it is possible that the distinct patterns observed for patients and typical speakers may be accounted for, in part, by differences in the degree of jaw involvement. Though trismus, or reduced jaw mobility, commonly results from oral cancer treatment (Watters *et al.*, 2019), bearing in mind task-dynamics and the synergistic relationships of the jaw and other oral articulators, patients may be expected to use larger jaw movements to compensate for reduced lingual mobility. Preliminary analyses of correlation between constriction degree and jaw height in the present data, however, suggest that the patient data does not pattern differently than the typical data in this respect.

Further, it is possible that specific aspects of vocal tract morphology (e.g., degree of concavity of the palate, length of the pharynx) contribute to the PC patterns exhibited, independently of any differences in lingual function due to glossectomy, among speakers. Future research focused on carrying out PC analyses on a large sample of typical individuals with a variety of distinct vocal tract features could

shed light on any systematic relationships that may exist between these morphological features and the PC patterns exhibited. Knowledge of these relationships could then be integrated into future applications of PC analysis to vocal tract cross-distance data from patients.

Furthermore, it must be acknowledged that the present investigation examines complexity of vocal tract shaping in only the midsagittal plane. Especially given that carcinomas oftentimes affect the tongue asymmetrically, it is possible that the degree of complexity measured in the midsagittal plane may vary appreciably from that measured in other sagittal slices [as consistent with Stone *et al.* (2014a), revealing different movement patterns between tumor-associated slices and slices in regions not affected by the tumor], as measured in other planes (e.g., coronal), and as measured when three-dimensional movement in the entire vocal tract is analyzed. Future directions of this work would benefit from leveraging recent developments in multi-planar and three-dimensional rtMRI data acquisition techniques (Kim *et al.*, 2009; Lim *et al.*, 2018) to provide a more holistic characterization of the complexity of vocal tract shaping in patients and typical speakers. Data of this type would be especially useful in identifying ways in which patterns exhibited by patients with lateral asymmetries (such as patient M5, who underwent unilateral resection) differ subtly from those produced by typical speakers.

Last, the preliminary results of this study, suggesting that patients exhibit less complex vocal tract shaping in the midsagittal plane than do typical speakers, have clinical implications. Reduced complexity of vocal tract shaping likely results, in large part, from patients' difficulty in differentially controlling distinct parts of the tongue, which is most likely to impact the production of speech segments involving multiple, synchronously produced lingual gestures (e.g., /l/, /ɹ/) as well as intersegmental co-articulation in natural speech. As such, it is likely that the most functional and generalizable post-surgical therapeutic exercises for patients would target complex lingual segments, as well as sequences requiring temporally overlapping tongue movements to form constrictions in diverse vocal tract regions, as occurs in natural speech production, rather than exercises targeting unidirectional lingual movement at a single point in time. The quantification of parameters indexing the complexity of vocal tract shaping using principal component analyses utilized in this study can be used also at various points post-operatively, during the course of treatment. In this way, the efficacy of diverse therapy methods in increasing the complexity of vocal tract shaping can be quantified and compared.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health Grants Nos. DC007124-01, DC008780-05, and DC009975-01. The authors are grateful to the participants and their families for their willingness to take part in this research.

TABLE III. Average speech rate for typical speakers and patients.

Average speech rate for typical speakers and patients	
TYPICAL SPEAKERS	
FT1	4.2 syl./s
MT1	4.3 syl./s
Average speech rate	4.25 syl./s
PATIENTS	
F1	2.9 syl./s
M1	2.2 syl./s
M2	3.2 syl./s
M3	2.9 syl./s
M4	3.7 syl./s
M5	3.7 syl./s
Average speech rate	3.1 syl./s

APPENDIX

Phrases read by participants.

TIMIT Sentences:

“She had your dark suit in greasy wash water all year.”

“Don’t ask me to carry an oily rag like that.”

Rainbow Passage Excerpt:

“When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arc with its path high above and its two ends apparently beyond the horizon.”

¹See the supplementary material <https://www.scitation.org/doi/suppl/10.1121/10.0004789> for tables of percentage of variability explained by components required to account for 99% of the variability in the data for all speakers.

Bachher, G. K., Dholam, K., and Pai, P. S. (2002). “Effective rehabilitation after partial glossectomy,” *Ind. J. Otolaryngol. Head Neck Surg.* **54**(1), 39–43.

Bilodeau-Mercure, M., and Tremblay, P. (2016). “Age differences in sequential speech production: Articulatory and physiological factors,” *J. Am. Geriatrics Soc.* **64**(11), e177–e182.

Bresch, E., Nielsen, J., Nayak, K., and Narayanan, S. (2006). “Synchronized and noise-robust audio recordings during realtime magnetic resonance imaging scans,” *J. Acoust. Soc. Am.* **120**(4), 1791–1794.

Bressmann, T., Ackloo, E., Heng, C., and Irish, J. C. (2007). “Quantitative three-dimensional ultrasound imaging of partially resected tongues,” *Otolaryngol. Head Neck Surg.* **136**(5), 799–805.

Bressmann, T., Uy, C., and Irish, J. C. (2005). “Analysing normal and partial glossectomy tongues using ultrasound,” *Clin. Ling. Phon.* **19**(1), 35–52.

Chandrasekara, R., Rao, A., Sanchez-Ortiz, G. I., Mohiaddin, R. H., and Rueckert, D. (2003). “Construction of a statistical model for cardiac motion analysis using nonrigid image registration,” in *Biennial International Conference on Information Processing in Medical Imaging* (Springer, Berlin, Heidelberg), pp. 599–610.

Daffertshofer, A., Lamoth, C. J., Meijer, O. G., and Beek, P. J. (2004). “PCA in studying coordination and variability: A tutorial,” *Clin. Biomech.* **19**, 415–428.

Fletcher, S. G. (1988). “Speech production following partial glossectomy,” *J. Speech Hear. Disorders* **53**, 232–238.

Georgian, D., Logemann, J. A., and Fischer, H. B. (1982). “Compensatory articulation patterns of a surgically treated oral cancer patient,” *J. Speech Hear. Disorders* **47**, 154–159.

Harshman, R., Ladefoged, P., and Goldstein, L. (1977). “Factor analysis of tongue shape,” *J. Acoust. Soc. Am.* **62**, 693–707.

Imai, S., and Michi, K. (1992). “Articulatory function after resection of the tongue floor of the mouth: Palatometric and perceptual evaluation,” *J. Speech Lang. Hear. Res.* **35**, 68–78.

Iskarous, K. (2005). “Patterns of tongue movement,” *J. Phon.* **33**, 363–381.

Kaipa, R., Robb, M. P., O’Biernie, G. A., and Allison, R. S. (2012). “Recovery of speech following total glossectomy: An acoustic and perceptual appraisal,” *Int. J. Speech-Lang. Pathol.* **14**(1), 24–34.

Kim, J., Kumar, N., Lee, S., and Narayanan, S. S. (2014). “Enhanced airway-tissue boundary segmentation for real-time magnetic resonance imaging data,” in *Proceedings of the International Seminar on Speech Production*, Cologne, Germany.

Kim, Y.-C., Narayanan, S., and Nayak, K. (2009). “Accelerated 3D upper airway MRI using compressed sensing,” *Magn. Reson. Med.* **61**(6), 1434–1440.

Koppetsch, S., and Dahlmeier, K. (2003). “Phonetic analysis of functional disorders of vocal articulation in cases of intra-oral carcinoma—A pre- and postoperative longterm study,” in *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, pp. 3249–3252.

Lim, Y., Lingala, S., Narayanan, S., and Nayak, K. (2018). “Dynamic off-resonance correction for spiral real-time MRI of speech,” *Magn. Reson. Med.* **81**, 234–246.

Logemann, J. A., Pauloski, B. R., Rademaker, A. W., McConnel, F. M., Heiser, M. A., Cardinale, S., Johnson, J., and Baker, T. (1993). “Speech

and swallow function after tonsil/base of tongue resection with primary closure,” *J. Speech Hear. Res.* **36**, 918–926.

Maeda, S. (1990). “Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model,” in *Speech Production and Speech Modeling*, edited by W. J. Hardcastle and A. Marchal (Kluwer Academic, Dordrecht), pp. 131–149.

McMicken, B., Von Berg, S., and Iskarous, K. (2012). “Acoustic and perceptual description of vowels in a speaker with congenital aglossia,” *Commun. Disord. Quart.* **34**(1), 38–46.

Michi, K., Imai, S., Yamashita, Y., and Suzuki, S. (1989). “Improvement of speech intelligibility by a secondary operation to mobilize the tongue after glossectomy,” *J. Cranio Maxillofacial Surg.* **17**, 162–166.

Michiwaki, Y., Schmelzeisen, R., Hacki, T., and Michi, K. (1993). “Functional effects of a free jejunum flap used for reconstruction in the oropharyngeal region,” *J. Craniomaxillofacial Surg.* **21**, 153–156.

Morrish, L. (1984). “Compensatory vowel articulation of the glossectomee: Acoustic and videofluoroscopic evidence,” *British J. Disorders Commun.* **19**(2), 125–134.

Oral Cancer Foundation (2021). “Oral cancer facts,” <http://oralcancerfoundation.org/facts/> (Last viewed July 18, 2017).

Ormonet, D., Black, M. J., Hastie, T., and Kjellström, H. (2005). “Representing cyclic human motion using functional analysis,” *Image Vision Comput.* **23**, 1264–1276.

Proctor, M. I., Bone, D., Katsamanis, N., and Narayanan, S. (2010). “Rapid semi-automatic segmentation of real-time magnetic resonance images for parametric vocal tract analysis,” in *Proceedings of the 10th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, Makuhari, Japan, pp. 1576–1579.

Savariaux, C., Perrier, P., Pape, D., and Jacques, L. (2001). “Speech production after glossectomy and reconstructive lingual surgery: A longitudinal study,” *The 2nd International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications*, Italy.

Shin, Y. S., Koh, Y. W., Kim, S. H., Jeong, J. H., Ahn, S., Hong, H. J., and Choi, E. C. (2012). “Radiotherapy deteriorates postoperative functional outcome after partial glossectomy with free flap reconstruction,” *J. Oral Maxillofacial Surg.* **70**(1), 216–220.

Stone, M., Epstein, M. A., and Iskarous, K. (2004). “Functional segments in tongue movement,” *Clin. Ling. Phon.* **18**(6-8), 507–521.

Stone, M., Langguth, J., Woo, J., Chen, H., and Prince, J. (2014a). “Tongue motion patterns in post-glossectomy and typical speakers: A Principal Components Analysis,” *J. Speech Lang. Hear. Res.* **57**(3), 707–717.

Stone, M., Liu, X., Chen, H., and Prince, J. L. (2010). “A preliminary application of principal components and cluster analysis to internal tongue deformation patterns,” *Comput. Methods Biomech. Biomed. Eng.* **13**, 493–503.

Stone, M., Woo, J., Zhuo, J., Chen, H., and Prince, J. L. (2014b). “Patterns of variance in /s/ during normal and glossectomy speech,” *Comput. Meth. Biomech. Biomed. Eng.: Imag. Visual.* **2**(4), 197–207.

Suma, H. N., and Murali, S. (2007). “Principal Component Analysis for analysis and classification of fMRI activation maps,” *Int. J. Comput. Sci. Netw. Security* **7**(11), 235–242.

Tabassian, M., Alessandrini, M., De Marchi, L., Masetti, G., Cauwenberghs, N., Kouznetsova, T., and D’hooge, J. (2015). “Principal component analysis for the classification of cardiac motion abnormalities based on echocardiographic strain and strain rate imaging,” in *Functional Imaging and Modeling of the Heart*, edited by H. van Assen, P. Bovendeerd, and T. Delhaas (Springer, Berlin).

Wang, X., O’Dwyer, N., and Halaki, M. (2013). “A review on the coordinative structure of human walking and the application of principal component analysis,” *Neural Regen. Res.* **8**(7), 662–670.

Watters, A. L., Cope, S., Keller, M. N., Padilla, M., and Enciso, R. (2019). “Prevalence of trismus in patients with head and neck cancer: A systematic review with meta-analysis,” *Head Neck* **41**(9), 3408–3421.

Wrench, A. A., and William, H. J. (2000). “A multichannel articulatory database and its application for automatic speech recognition,” in *5th Seminar on Speech Production: Models and Data, Bavaria*, pp. 305–308.

Xing, F., Stone, M., Goldsmith, T., Prince, J. L., El Fakhri, G., and Woo, J. (2019). “Atlas-based tongue muscle correlation analysis from tagged and high-resolution magnetic resonance imaging,” *J. Speech Lang. Hear. Res.* **62**(7), 2258–2269.

Zhou, X., Stone, M., and Carol, E.-W. (2011). “A Comparative Acoustic Study on Speech of Glossectomy Patients and Normal Subjects,” *Proceedings of Interspeech*, Florence, Italy.