



**HAL**  
open science

## Sensorimotor adaptation across the speech production workspace in response to a palatal perturbation

Guillaume Barbier, Shari R Baum, Lucie Ménard, Douglas M Shiller

► **To cite this version:**

Guillaume Barbier, Shari R Baum, Lucie Ménard, Douglas M Shiller. Sensorimotor adaptation across the speech production workspace in response to a palatal perturbation. *Journal of the Acoustical Society of America*, 2020, 147, pp.1163 - 1178. 10.1121/10.0000672 . hal-03919657

**HAL Id: hal-03919657**

**<https://hal.science/hal-03919657v1>**

Submitted on 3 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Sensorimotor adaptation across the speech production workspace in response to a palatal perturbation

Guillaume Barbier,<sup>1</sup> Shari R. Baum,<sup>2</sup> Lucie Ménard,<sup>3</sup> and Douglas M. Shiller<sup>1,a)</sup>

<sup>1</sup>*École d'orthophonie et d'audiologie, Université de Montréal, Case Postale 6128, Succursale Centre-Ville, Montréal, Québec, H3C 3J7, Canada*

<sup>2</sup>*School of Communication Sciences and Disorders, McGill University, 2001 McGill College Avenue, Suite 800, Montréal, Québec, H3A 1G1, Canada*

<sup>3</sup>*Département de linguistique, Université du Québec à Montréal, Case Postale 8888, Succursale Centre-Ville, Montréal, Québec, H3C 3P8, Canada*

### ABSTRACT:

Talkers have been shown to adapt the production of multiple vowel sounds simultaneously in response to altered auditory feedback. The present study extends this work by exploring the adaptation of speech production to a physical alteration of the vocal tract involving a palatal prosthesis that impacts both somatosensory and auditory feedback during the production of a range of consonants and vowels. Acoustic and kinematic measures of the tongue were used to examine the impact of the physical perturbation across the various speech sounds, and to assess learned changes following 20 min of speech practice involving the production of complex, variable sentences. As in prior studies, acoustic analyses showed perturbation and adaptation effects primarily for sounds directly involving interaction with the palate. Analyses of tongue kinematics, however, revealed systematic, robust effects of the perturbation and subsequent motor learning across the full range of speech sounds. The results indicate that speakers are able to reconfigure oral motor patterns during the production of multiple speech sounds spanning the articulatory workspace following a physical alteration of the vocal tract. © 2020 Acoustical Society of America. <https://doi.org/10.1121/10.0000672>

(Received 21 June 2019; revised 14 January 2020; accepted 15 January 2020; published online 19 February 2020)

[Editor: Susanne Fuchs]

Pages: 1163–1178

### I. INTRODUCTION

The speech of adult talkers is generally characterized by a high degree of spatiotemporal precision in the control of oral articulators such as the tongue, including the achievement of precise spatial targets within very tight temporal constraints (Smith and Zelaznik, 2004). Another key characteristic of speech motor control in mature talkers, as in other motor domains, is the capacity for adaptation to novel conditions, i.e., sensorimotor plasticity. The capacity for speech motor adaptation has been examined in numerous experimental studies involving real-time alterations of sensory feedback during speech production. Such manipulations have included perturbations focusing on auditory feedback (Houde and Jordan, 1998; Purcell and Munhall, 2006; Cai *et al.*, 2011) and somatosensory feedback (Tremblay *et al.*, 2003; Nasir and Ostry, 2006) in relative isolation, in addition to physical manipulations that impact both auditory and somatosensory modalities simultaneously (Kelso *et al.*, 1984; Savariaux *et al.*, 1995; Baum and McFarland, 1997; Jones and Munhall, 2003). These studies have consistently demonstrated that adult talkers are capable of adapting their speech motor strategies to altered sensorimotor conditions,<sup>1</sup> and have highlighted the role of auditory and somatosensory feedback and prediction in speech motor learning and control (see Perkell, 2012, for review).

While studies of speech adaptation have played a central role in the development and evaluation of current models of speech production (Tourville and Guenther, 2011; Houde and Nagarajan, 2011; Ramanarayanan *et al.*, 2016; Parrell *et al.*, 2019), the vast majority of these studies have examined motor adaptation in very narrow contexts within any given experiment (e.g., one vowel within a small set of words, or a limited number of speech sounds sharing major characteristics within a handful of phonetic contexts or acoustically similar sentences). While providing carefully controlled demonstrations of motor learning, these studies have involved speech tasks that differ considerably from speech production under natural conditions, where multiple sounds spanning the articulatory workspace (i.e., covering the range of high, low, anterior, and posterior tongue articulatory positions) are produced in the context of highly variable, fluid sentences.

In contrast to these investigations, a recent study by Lametti *et al.* (2018) investigated the capacity of talkers to reorganize the control of speech motor patterns across the entire articulatory vowel space following a perturbation of auditory feedback affecting all vowel sounds during the production of complex, variable sentences. In this study, a systematic shift in the first two vowel resonant (formant) frequencies was introduced that systematically altered all vowels during the production of 50 different English sentences. Following a brief period of intensive speech practice under altered feedback conditions (200 sentences, ~20 min),

<sup>a)</sup>Electronic mail: douglas.shiller@umontreal.ca

sounds across the workspace showed reliable adaptation effects that countered the auditory perturbation, as well as robust transfer to isolated words containing the same vowel sounds. The result demonstrated that despite the considerable sensorimotor complexity of connected speech, auditory feedback is continually used to adapt oral motor patterns across speech sounds in response to perceived deviations from auditory sensory targets.

While auditory input plays a central role in the learning and control of speech motor behaviors, there is considerable experimental work showing that speech production relies critically on both auditory and somatosensory feedback in order to achieve targets in both sensory domains (Tremblay *et al.*, 2003; Nasir and Ostry, 2006; Lametti *et al.*, 2012). A variety of structural and functional manipulations of the vocal apparatus during speech production have been used to explore this, including jaw fixation using a bite-block held between the teeth (e.g., Gay *et al.*, 1981), holding a tube between the lips (e.g., Savariaux *et al.*, 1995), and wearing an intra-oral prosthesis that lengthens the upper incisors (Jones and Munhall, 2003) or alters the shape of the hard palate (e.g., Hamlet *et al.*, 1976). These studies have consistently shown that talkers are capable of adapting to such physical perturbations, leading to novel articulatory strategies that improve the quality of the acoustic output. As in studies focusing exclusively on altered auditory-feedback, however, the vast majority of studies examining the effects of physical perturbations have been somewhat narrow in scope. For example, the majority of bite-block studies and lip-tube studies have involved the production of a small number of vowels, either in isolation or in a limited number of word-like contexts (although, see McFarland *et al.*, 1996). Similarly, studies involving dental prostheses have focused on sibilant consonants (/s/) practiced in a simple monosyllabic word-like context (Jones and Munhall, 2003).

The majority of studies examining speech adaptation to experimental changes in palatal shape have also been narrow in scope, typically focusing on consonants involving direct contact with the physically altered alveolar region in simple phonetic contexts (Hamlet *et al.*, 1978; Baum and McFarland, 1997; Baum and McFarland, 2000; Honda *et al.*, 2002; Aasland *et al.*, 2006; Brunner *et al.*, 2011). However, a number of studies have examined the effects of this structural modification beyond the palatal region, including stop consonants (e.g., /k/; Brunner, 2009; McFarland *et al.*, 1996; Thibeault *et al.*, 2011) and vowels (Brunner *et al.*, 2007; Brunner, 2009; Hamlet *et al.*, 1976; Garber *et al.*, 1980; McFarland *et al.*, 1996). These studies have generally observed only limited acoustic effects of the palatal perturbation beyond alveolar sounds (e.g., /s/ or /t/), often noting little systematic acoustic change in the production of vowels (Brunner, 2009; Garber *et al.*, 1980; McFarland *et al.*, 1996; although see Hamlet *et al.*, 1976), and only negligible effects on non-alveolar stop consonants (/k/ or /p/; McFarland *et al.*, 1996; Thibeault *et al.*, 2011). While the acoustic effects of this manipulation have appeared mostly restricted to sounds directly involving

interactions with the palate, *kinematic* measures of tongue movement patterns reported in studies by Brunner *et al.* (2009) and Thibeault *et al.* (2011) have suggested that such palatal perturbations may, in fact, impact articulatory speech movements beyond the palatal region (e.g., low/back vowels, or the velar stop consonant /k/). This finding, coupled with the large degree of inter-subject variability in the patterns of oral perturbation and adaptation noted by many authors (Baum and McFarland, 2000; Aasland *et al.*, 2006, Brunner, 2009; Thibeault *et al.*, 2011), leaves open the question of whether talkers are capable of systematically adapting speech motor patterns for sounds produced across the articulatory workspace to a physical structural perturbation of the vocal apparatus.

In the present study, we address this question by examining, through a combination of acoustic and tongue kinematic measures, the manner in which talkers adapt oral motor patterns to a physical alteration of palatal shape during the production of multiple speech sounds across the articulatory workspace, including a variety of consonants and vowels. Specifically, we address the following questions:

- (1) To what degree does an experimental alteration of palatal shape impact the *acoustic* outcomes related to the production of phonemes across the articulatory workspace, and to what degree do talkers show compensatory changes in speech acoustics across these different sounds following 20 min of intensive speech practice with the palate in place? Based on prior studies, we predict that acoustic changes related to the perturbation, and in response to speech practice, will be observed selectively in sounds that interact directly with the palatal perturbation.
- (2) To what degree does the palatal prosthesis perturb tongue *kinematic* patterns related to a range of speech sounds across the workspace, and to what extent do talkers show sensorimotor adaptation in motor control of the tongue for these sounds following an intensive 20-min practice period? This is a question that, we feel, has not been clearly answered in prior studies, owing to the highly variable nature of the kinematic effects of the perturbation and adaptation response among participants and phonetic contexts. In the present study, we approach this question in two steps: First, by examining the *amplitude* of tongue kinematic changes associated with the presence of the palate and following speech practice, it is possible to determine whether or not the alteration in palatal shape induces perturbation and training-related changes in tongue position across the various speech sounds. Second, by examining the *directional* nature of these kinematic changes, it is possible to address the question of whether the effects of practice reflect learned changes in the feed-forward control of the tongue (i.e., sensorimotor adaptation). Based upon the recent demonstration of simultaneous adaptation in multiple vowels in response to a manipulation of auditory feedback (Lametti *et al.*, 2018), we predict that talkers

in the present study will exhibit significant adaptation in speech motor patterns to a physical change in palate shape during the production of phonemes that span the production workspace, as revealed by both the amplitude and directions of the perturbation and training effects.

The examination of the directional nature of kinematic changes in the current study rests critically on a consideration of the *learning after-effect*. In current models of the sensorimotor control of speech production, oral movements are controlled by a combination of *feedback* and *feed-forward* mechanisms (see Parrell *et al.*, 2019, for review). In feedback control, sensory signals are compared with predicted sensory consequences in near-real-time, which results in the generation of corrective changes in production when a deviation is detected during a given movement. Feed-forward control, on the other hand, involves the use of predictive motor commands that are gradually established and fine-tuned on the basis of practice (i.e., motor learning). In studies of sensorimotor adaptation to alterations in sensory feedback, the presence of motor learning following a period of practice can be examined by suddenly removing the sensory perturbation (hence, eliminating the source of feedback-based error correction), and testing whether a systematic change in motor output persists. Such a persistent change, dubbed the learning after-effect, provides a window into the altered feed-forward (planned) motor command, and would be expected to show a systematic relationship with the changes observed during training while under perturbed sensory conditions. Numerous prior studies of sensorimotor adaptation in speech involving changes in auditory or somatosensory feedback have examined, and found, such learning after-effects (e.g., Houde and Jordan, 1998; Baum and McFarland, 1997; Shiller *et al.*, 2009; Villacorta *et al.*, 2007).

In the current study, we apply this rationale to a consideration of the directions of kinematic changes observed immediately following a period of speech practice, and the directions of the learning after-effects. We predict that irrespective of variation across subjects and phonemes in tongue kinematic changes associated with speech training, the directions of observed learning after-effects will be coherent with the directions of the preceding training effects, indicating that the talkers had indeed acquired new feed-forward speech motor commands in adapting to the palatal perturbation.

**II. METHODS**

Nine adult native speakers of North-American English (four female, five male, 19–30 years of age) with no history of hearing, language or speech disorder were tested. All procedures were approved by the Institutional Review Board of the Faculty of Medicine at McGill University.

**A. Procedures**

For each participant, a palatal prosthesis was custom-made by taking an alginate impression of the upper teeth and hard palate and producing a stone cast. A biocompatible

thermoplastic material (Temp Tabs, All Dental PRODX, Gig Harbor, WA) was molded to the alveolar region of the cast (immediately behind the upper teeth) producing a rigid prosthetic appliance with dimensions similar to those used in prior palatal perturbation studies (Baum and McFarland, 1997; Thibeault *et al.*, 2011), involving a 6 mm thickened region behind the upper incisors, tapering off over a 2 cm distance (Fig. 1). The appliance was held firmly in place by way of a thin layer of thermoplastic material that wrapped around the upper incisors.

Participants engaged in a series of four speech tests, each involving the production of eight different vowel-consonant-vowel pseudo-words containing one of the consonants /s/, /t/, /k/, and /p/ combined with either the high vowel /i/ or the low vowel /æ/ (e.g., /isi/, /ætæ/). Each pseudo-word was produced eight times in a randomized order (64 productions in total per test). Four such tests of speech production were conducted in sequence as follows: (1) prior to palatal prosthesis insertion (*Test-1*), (2) immediately after prosthesis insertion but prior to the period of speech practice with the palate in place (*Test-2*), (3) immediately following speech production practice with the palatal prosthesis in place (*Test-3*), and (4) immediately following removal of the prosthesis (*Test-4*; see Fig. 2). A period of speech practice with the prosthesis in place followed speech *Test-2*, during which participants read aloud a randomized sequence of sentences drawn from the Harvard Sentences (IEEE, 1969). A set of 47 different sentences was used, containing approximately balanced proportions of the four consonants included in the speech test. The 47 sentences were

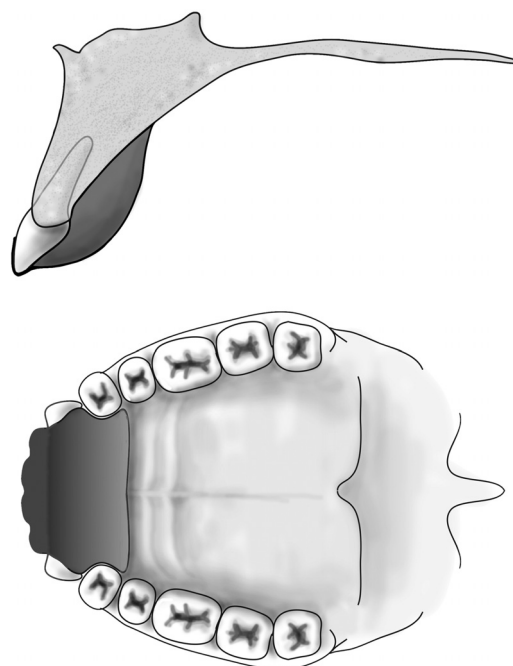


FIG. 1. Schematic showing the position of the palatal prosthesis (dark grey region) in the mouth. Top: Mid-sagittal view; Bottom: View from below the entire maxillary region, showing the palate and teeth. (Adapted from Baum and McFarland, 1997).

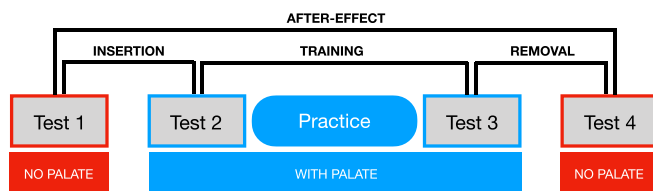


FIG. 2. (Color online) Schematic of the experimental sequence, showing the different speech tasks (middle), the sequence of insertion and removal of the palatal prosthesis (bottom), and the comparisons between tests that comprise the four experimental effects of interest (top).

read four times in randomized order, totaling 188 sentences (~20 min) of speech practice.

## B. Acoustic recording and analysis

The acoustic speech signal was digitally recorded (16-bit, 22 500 kHz) using a cardioid microphone (ME66, Sennheiser, Germany) positioned at a distance of one meter to reduce electromagnetic interference with the kinematic measurement system. Following prior studies of adaptation to palatal prostheses (Baum and McFarland, 1997; Aasland *et al.*, 2006; Thibeault *et al.*, 2011), acoustic analyses of the consonants focused on the first spectral moment (or *centroid* frequency), a primary acoustic cue for /s/ (Maniwa *et al.*, 2009), as well as a key contrastive property of stop consonant bursts (Forrest *et al.*, 1988; Johnson *et al.*, 2018). The centroid frequency was estimated over a 20 ms analysis window at the mid-point of production for the fricative /s/ and aligned with the burst onset for the stop consonants /t/, /k/, and /p/. For the vowels, the first and second formant frequency ( $F1$  and  $F2$ ) over a 40-ms window at the vowel mid-point was estimated using a Linear Predictive Coding based spectral analysis (MATLAB v.2018a, Mathworks, Natick, MA). For simplicity, the analysis of vowel formants focused solely on the syllable-initial position in each VCV production.

### 1. Statistical analyses

Statistical analyses were organized around four separate experimental effects of interest: (1) the *palate insertion* effect (Test-2 - Test-1), (2) *speech training* effect with the palate in place (Test-3 - Test-2), (3) the *palate removal* effect (Test-4 - Test-3), and (4) the learning *after-effect* (Test-4 - Test-1; see Fig. 1, top), using a linear mixed-effects modeling approach using R (R Core Team, 2019) with *lme4* (Bates *et al.*, 2019). For each of the four experimental effects, the significance of differences in the acoustic measures was tested by fitting the model

$$\text{Acoustic.Measure} \sim \text{VCV} * \text{Test} + (1 | \text{Participant}),$$

where *Acoustic.Measure* corresponds to formant frequency (for vowels) or centroid frequency (for consonants), *VCV* refers to the eight stimuli (/s/, /t/, /k/, and /p/ combined with /i/ and /æ/; with /isi/ as the reference level), *Test* corresponds to the two speech production tests comprising the experimental effect of interest (e.g., Test 2 vs Test 1 for the effect

of insertion), and (1 | Participant) represents the inclusion of random intercepts per participant. The statistical significance of fixed effects (including the two-way interaction of VCV and Test) was evaluated using the package *lmerTest* (v.3.1) in R, which provides analysis of variance (ANOVA)-style tables for linear mixed effects model fits via Satterthwaite's degrees of freedom method, allowing for the reporting of readily interpretable *F*-values, degrees-of-freedom, and *p*-values. *Post hoc* pairwise tests between levels of the fixed effects were carried out when necessary by computing estimated marginal means using the *emmeans* (v.1.4.1) in R and applying Tukey's correction for multiple comparisons.

## C. Kinematic measurement and analysis

Kinematic measurements were carried out using an electromagnetic articulograph (AG500, Carstens, Germany), with two sensors placed along the mid-sagittal axis of the tongue: one 1.5 cm from the apex (labelled tongue *blade*) and another located 4.5 cm from the apex (tongue *dorsum*). An additional sensor was affixed to the mandible at the lower incisors. Sensors affixed to the upper incisor, as well as to the surface of the skin at the nasion, left mastoid, and right mastoid, were used to establish a head-centered coordinate frame, allowing for the transformation of raw three-dimensional (3D) tongue sensor positions into positions relative to an origin located at the upper incisors, and with the horizontal axis registered to the occlusal plane. Positional data were sampled at 200 Hz. The kinematic analyses carried out in the current paper focused solely on positions of the two tongue markers in the mid-sagittal plane.

For the purposes of kinematic analysis, the position of the tongue sensors (horizontal, *x*, and vertical, *y*, position in the midsagittal plane) associated with the production of the vowels /i/ and /æ/ was selected at the temporal mid-point of production based on the acoustic signal. As the consonants were embedded in a vowel context, it was possible to select their positions based upon kinematic criteria. Following Thibeault *et al.* (2011), the tongue position for /s/ and /t/ was measured at the first zero-crossing in velocity along the vertical axis (i.e., peak tongue elevation) for the tongue blade sensor during the gesture from the syllable-initial vowel to the consonant. The position for /k/ in the low-vowel context (/aka/) was based on the same criteria, but using the tongue dorsum sensor. For /k/ in the high-vowel context (/iki/), the position was measured at the first zero-crossing in *horizontal* velocity (i.e., peak retraction) for the tongue dorsum sensor. Finally, /p/ was measured at the moment of maximal lip closure (minimum Euclidean distance between the upper and lower lip sensors).

As with the acoustic measures, statistical analyses were organized around the change in tongue position associated with each of the four key effects of interest: *palate insertion* (Test 2 – Test 1), *training* (Test 3 – Test 2), *palate removal* (Test 4 – Test 3), and the *learning after-effect* (Test 4 – Test 1). The analyses separately examined the *amplitude* of changes in tongue position and the *direction* of those changes

associated with the different speech sounds and phonetic contexts. In this way, it was possible to separately assess (1) whether a change in tongue position has occurred in each case (indicated by a significant change in position between a given pair of test conditions), and (2) the directional nature of any such positional changes, including whether those directions varied systematically among VCV conditions, experimental effects, and participants.

The *direction* of changes in tongue sensor position was determined on the basis of the average two-dimensional (2D) sensor position, within each participant, across the eight repeated productions of each VCV within each of the four speech tests. Specifically, a vector was computed for each VCV and experimental effect within each participant, characterizing the change in tongue position between the pair of speech tests associated with the experimental effects of interest (palate insertion, training, removal, and after-effect). For example, the direction of the change in tongue blade sensor position associated with the effect of palate insertion for the production of the vowel /æ/ would involve computing the angle of the vector whose origin lies at the mean ( $x, y$ ) position for the production of that sound during Speech Test 1 (pre-insertion), and whose direction points at the mean ( $x, y$ ) position for the production of that sound in Speech Test 2 [post-insertion; see Fig. 3(A)]. Recall that the horizontal ( $x$ ) axis is aligned with the participant's occlusal plane, with 0-degrees (graphically to the right) corresponding to a change towards the anterior of the mouth, and 90-degrees (upwards) corresponding to an increase in elevation. As such, the vectors are presented in a common coordinate system that is normalized for differences among participants and VCV contexts.

For the analysis of the *amplitude* of tongue position change associated with each experimental effect, the 2D spatial positions associated with individual productions were transformed into a one-dimensional (1D) positional representation that greatly simplified statistical testing using a univariate linear mixed-effects modelling approach. First, for each subject, VCV and experimental effect, an axis was defined on the basis of the mean direction vector characterizing the experimental effect of interest (the same vector computed in the directional analysis described above). The 2D ( $x, y$ ) positions of the tongue sensors associated with the participant's individual productions were then orthogonally projected onto this axis (dubbed the *Principal Effect Axis*), transforming them into a 1D representation of position along this axis of interest [see Figs. 3(B) and 3(C)].

### 1. Statistical analyses

The significance of differences in vector directions (angles) was examined using circular (or directional) statistics, a set of methods for evaluating observations representing vectors around a unit circle (Circular Statistics Toolbox for MATLAB, v.1.21; Berens, 2009). Such directional data necessitates an approach to mean and variance estimation that accounts for the "wrapped" nature of angles distributed

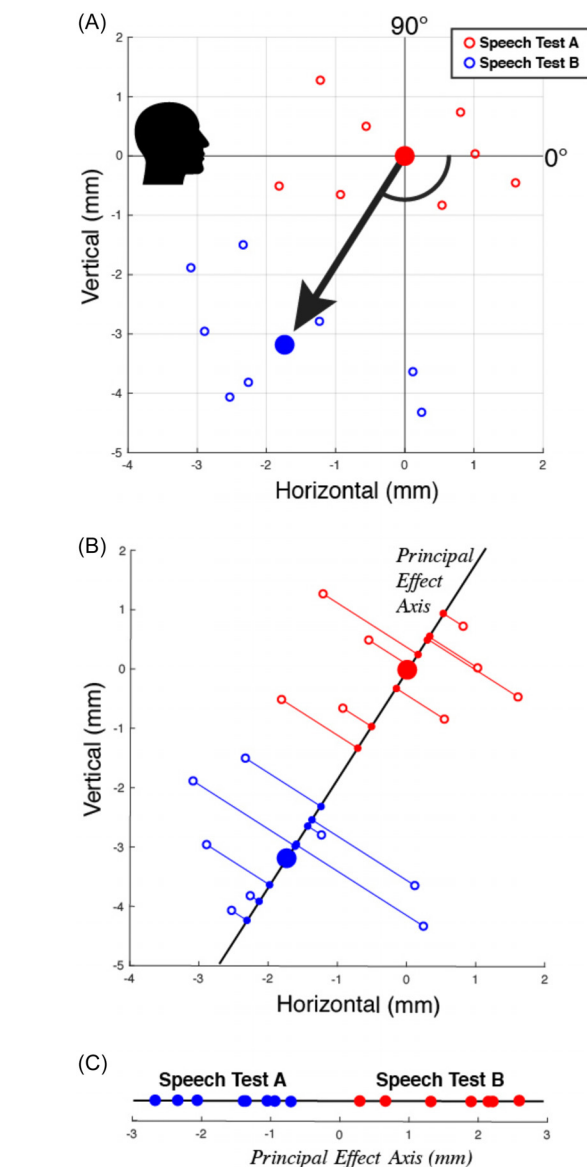


FIG. 3. (Color online) (A) Illustration of the procedure used to compute the vector angle characterizing an experimental effect of interest. (B) Illustration of the procedure for examining the *amplitude* of tongue position change associated with a given experimental effect. The 2D positions associated with individual productions are projected orthogonally onto the principal axis defining the experimental effect. (C) The resulting 1D positions along this axis can then be evaluated statistically (e.g., testing for an effect of speech test).

around a circle (where 0° = 360°). Differences in vector angles among phonemes and between experimental effects were examined using circular 1-way ANOVA (*Watson-Williams* test). Variation in vector angles among VCV contexts and participants was also evaluated using the *Rayleigh z* test, which assesses whether a set of vectors is non-uniformly distributed around a circle (i.e., characterized by a common mean direction).

For the analysis of the *amplitude* of changes in tongue position, differences in sensor position between the two speech tests characterizing each experimental effect was tested statistically by fitting a linear mixed-effects model of the following form:

$$\text{Kinematic.Measure} \sim \text{VCV} * \text{Sound} * \text{Test} + (1 | \text{Participant}),$$

where *Kinematic.Measure* corresponds to sensor position along the Principal Effect Axis, *VCV* refers to the eight different VCV contexts (with /isi/ as the reference level), *Sound* refers to the sound within the syllable (syllable-initial vowel or consonant), *Test* corresponds to the two speech production tests associated with the experimental effect of interest, and (1 | Participant) represents the inclusion of random intercepts per participant. The statistical significance of fixed effects (including the 2-way and 3-way interactions between fixed factors) was evaluated using *F*-tests with Satterthwaite estimates of degrees-of-freedom.

All data and code related to the statistical analyses reported in the current manuscript are available for download from the Open Science Framework repository.<sup>2</sup> Full text output of the linear mixed effects regression models are available as a supplement to this paper.<sup>3</sup>

### III. RESULTS

#### A. Acoustics – Consonants

##### 1. Effects of palate insertion and removal

The average changes in centroid frequency during the production of the four consonants observed following palate insertion and removal are shown in Fig. 4(A). Immediately after insertion of the palate (blue bars), the centroid frequency can be seen to decrease for production of /s/, /t/, and to a lesser degree /k/, in both vowel contexts. Following removal of the palate (red bars), a change in centroid frequency in the opposite direction (increasing) can be observed for all sounds. The difference between the two

production tests associated with palate insertion (Test 1 vs Test 2) and removal (Test 3 vs Test 4), as well as differences among the eight VCV conditions, were examined by fitting a linear mixed-effects model for each experimental effect. For the effect of insertion, a significant fit of speech test was found [ $F(1,1111) = 202.89, p < 0.001$ ], as well as for VCV [ $F(7,1111) = 277.24, p < 0.001$ ]. The two-way interaction between VCV and speech test was also significant [ $F(7,1111) = 18.03, p < 0.001$ ]. *Post hoc* pairwise comparisons between the two speech tests were carried out separately for each VCV stimulus, revealing a significant effect of palate insertion for /s/, /t/, and /k/ in the /i/ context, and for /s/ and /t/ in the /æ/ context ( $p < 0.01$ ). All remaining comparisons were not significant ( $p > 0.05$ ).

The effect of palate removal (Test 3 vs Test 4) showed the same pattern of results, with highly significant fits for the effect of speech test [ $F(1,1109) = 155.85, p < 0.001$ ], as well as VCV [ $F(7,1109) = 321.67, p < 0.001$ ], and a significant interaction [ $F(7,1109) = 16.75, p < 0.001$ ]. *Post hoc* tests of the effect of speech test within each VCV showed a significant difference for /s/, /t/, and /k/ in the /i/ context, and for /s/ and /t/ in the /æ/ context ( $p < 0.01$ ). All remaining comparisons were not significant ( $p > 0.05$ ).

##### 2. Effects of speech training and the learning after-effect

Figure 4(B) shows the average change in centroid frequency following 20-min of speech practice involving the production of variable sentences (green bars), and the learning after-effect (yellow bars). These changes are relatively small in amplitude compared to the effects of palate insertion and removal (particularly for /s/ and /t/).

The differences between the two speech tests associated with speech practice (Test 2 vs Test 3) and the learning

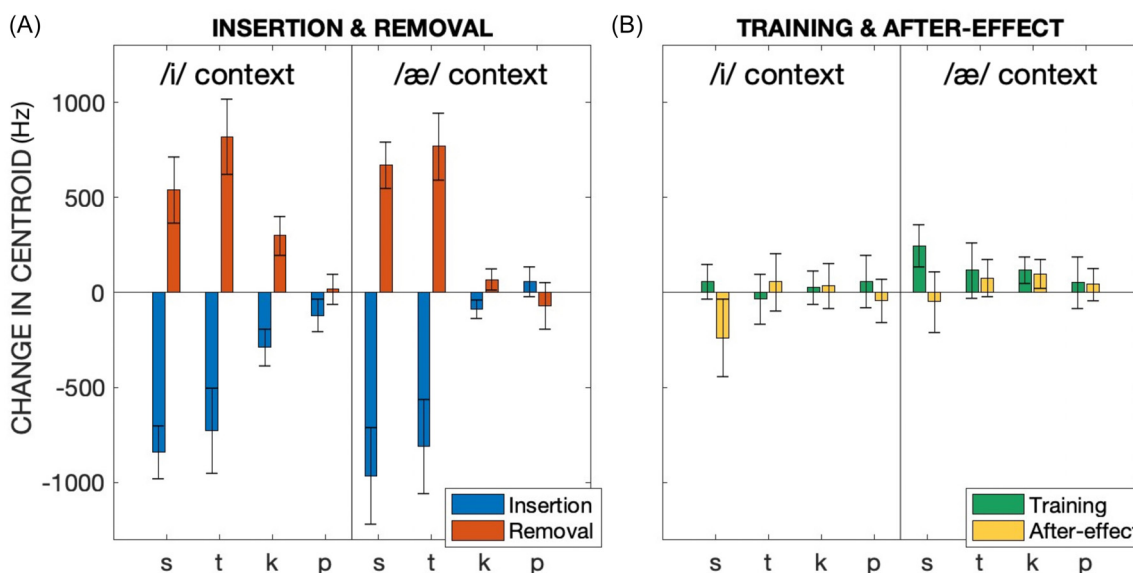


FIG. 4. (Color online) (A) Changes in centroid frequency during the production of the four consonants associated with the effects of palate insertion (blue) and removal (red), within the two vowel contexts. (B) Changes in centroid frequency during consonant production associated with the effects of training (green) and the learning after-effect (yellow). Error bars show  $\pm 1$  standard error of the mean.

after-effect (Test 1 vs Test 4) were examined in combination with differences in VCV condition by fitting a linear mixed-effects model for each experimental effect. For the model testing the effect of speech practice, the fit of VCV was significant [ $F(7,1102) = 178.61, p < 0.001$ ], as was the effect of speech test [ $F(1,1102) = 6.41, p < 0.05$ ]; however, the interaction was not [ $F(7,1102) = 0.64, p = 0.72$ ]. Because of the specific focus on the fricative /s/ in prior studies using palatal perturbations, this sound was directly examined for an effect of speech training. A *post hoc* comparison between speech tests for /s/ in the two vowel contexts revealed a significant difference in the /æ/ context ( $p < 0.05$ ), but not in the /i/ context ( $p = 0.67$ ).

For the learning after-effect (Test 1 vs Test 4), the fit of VCV was significant [ $F(7,1118) = 435.48, p < 0.001$ ], but not the effect of speech test [ $F(1,1118) = 0.090, p = 0.76$ ]. The interaction was also not statistically significant [ $F(7,1118) = 1.35, p = 0.22$ ].

In summary, as shown in prior studies, a reliable change in centroid frequency was observed following palate insertion and removal for the alveolar consonants /s/ and /t/. Though smaller in amplitude, an effect was also observed for the velar consonant /k/. Further consistent with prior studies, a change in acoustics was observed following speech training for /s/, though only in the low vowel context. No other consonant showed a significant training effect or learning after-effect.

## B. Acoustics—Vowels

### 1. Effects of palate insertion and removal

The mean changes in first and second formant frequency for the vowels associated with palate insertion and removal are shown in Fig. 5(A). For both /i/ and /æ/, a change can be

observed principally in  $F2$ , which shows a systematic reduction following insertion and an increase following removal.

A linear mixed-effects model was used to test for differences between the two speech tests associated with the experimental effects of palate insertion and removal, in conjunction with the eight different VCV stimuli, separately for  $F1$  and  $F2$ .

For the effect of palate insertion (Test 1 vs Test 2) on  $F1$  frequency, the effect of VCV was statistically significant [ $F(7,1111) = 1193.32, p < 0.001$ ]; however, the fit of speech test was not significant [ $F(1,1111) = 2.01, p = 0.16$ ], nor was the interaction between VCV and speech test [ $F(7,1111) = 0.84, p = 0.55$ ]. The effect of palate removal (Test 3 vs Test 4) on  $F1$  frequency followed the same pattern as insertion, with a significant effect of VCV [ $F(7,1109) = 1228.10, p < 0.001$ ], and no significant effect of speech test [ $F(1,1109) = 0.45, p = 0.50$ ] or interaction [ $F(7,1109) = 1.24, p = 0.28$ ].

In contrast with  $F1$ , the effect of palate insertion on  $F2$  values revealed a significant effect of speech test [ $F(1,1102) = 25.81, p < 0.001$ ] and of VCV [ $F(7,1102) = 2067.61, p < 0.001$ ], as well as a significant interaction [ $F(7,1070) = 2.66, p < 0.01$ ]. *Post hoc* tests showed significant differences between the two speech tests for the vowel /i/ in the contexts /s/ and /t/ ( $p < 0.001$ ), and for the vowel /æ/ in the context /s/ ( $p < 0.01$ ). All remaining comparisons were not significant ( $p > 0.05$ ). The effect of palate removal on  $F2$  values similarly revealed a significant effect of speech test [ $F(1,1092) = 32.46, p < 0.001$ ] and VCV [ $F(7,1092) = 1533.40, p < 0.001$ ] as well as a significant interaction [ $F(7,1053) = 4.19, p < 0.001$ ]. *Post hoc* tests showed a significant difference between the two speech tests for the vowel /i/ in all four consonant contexts ( $p < 0.001$ ), and for the vowel /æ/ in the context /t/ ( $p < 0.01$ ). All remaining comparisons were not significant ( $p > 0.05$ ).

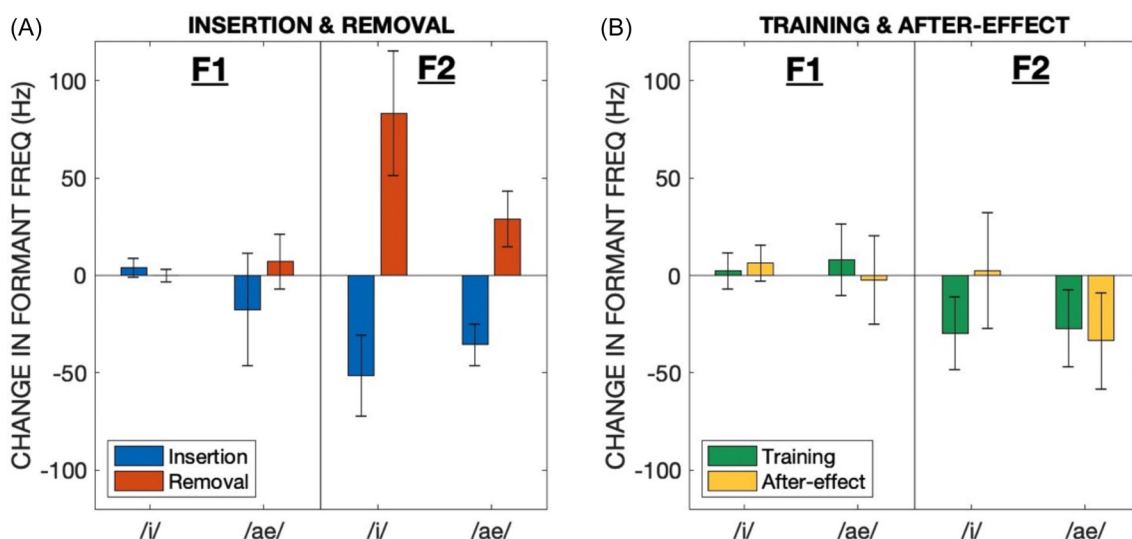


FIG. 5. (Color online) (A) Acoustic changes in vowels (averaged across the four consonant contexts) associated with palate insertion (blue) and removal (red).  $F1$  and  $F2$  are shown in the left and right half of the panel, respectively. (B) Acoustic changes in  $F1$  and  $F2$  associated with the training (green) and after-effect (yellow). Error bars show  $\pm 1$  standard error of the mean.



## 2. Effects of speech training and the learning after-effect

The change in vowel  $F1$  and  $F2$  associated with speech training and the learning after-effect are shown in Fig. 5(B). Modest changes can be observed in  $F2$ , with smaller changes in  $F1$ , mirroring the larger effects in  $F2$  associated with palate insertion and removal.

For the model testing the effect of training on  $F1$ , the effect of VCV was significant [ $F(7,1102) = 1227.85, p < 0.001$ ]; however, the fit of speech test was not significant [ $F(1,1102) = 1.10, p = 0.30$ ], nor was the interaction [ $F(7,1102) = 1.60, p = 0.13$ ]. For the learning after-effect, the effect of VCV was significant [ $F(7,1118) = 1206.41, p < 0.001$ ], while the effect of speech test [ $F(1,1118) = 0.11, p = 0.73$ ] and the interaction [ $F(7,1118) = 0.18, p = 0.99$ ] were not.

For the model testing the effect of training on  $F2$ , the principal effect of speech test was significant [ $F(1,1093) = 9.13, p < 0.01$ ], as was the effect of VCV [ $F(7,1093) = 1910.85, p < 0.001$ ]. The interaction was not significant [ $F(7,1093) = 0.37, p = 0.92$ ]. For the after-effect, the effect of VCV was significant [ $F(7,1101) = 1644.67, p < 0.001$ ]; however, no significant effect of speech test [ $F(1,1101) = 1.67, p = 0.20$ ] and no interaction [ $F(7,1101) = 0.78, p = 0.61$ ] were found.

In summary, an effect of palate insertion and removal was found to have an impact on the production of both the high and low vowels, in particular in  $F2$  frequency. In addition, a statistically reliable change in  $F2$  frequency was observed following speech training with the palatal prosthesis in place; however, no significant learning after-effect was found.

## C. Kinematics—Amplitude

The average amplitude of changes in tongue sensor position associated with the four experimental effects of interest is shown for all consonants and vowels in Fig. 6. A systematic change in tongue position can be observed in both the blade and dorsum sensors across all sounds for each of the four effects. Note that the analysis of amplitude related to the training and after-effects omitted the data of a single outlier participant (s7) who showed a magnitude of training that was considerably *larger* than all other subjects—by a factor of more than two—across nearly all sounds and contexts. We chose to exclude this participant from the analysis of amplitude in order to avoid exaggerating the group-average magnitude of the training effect, as well as to avoid inflating the estimate of within-group variance. For the remaining analyses of changes in tongue kinematics, all nine participants were included.

The change in spatial position of each tongue sensor—associated with each of the four experimental effects of interest—was examined using a linear mixed-effects model with speech test, VCV, and speech sound (vowel vs consonant) as fixed factors (see Sec. II). For the tongue *dorsum* sensor, the analysis revealed a statistically significant fit for the effect of speech test associated with palate insertion

[Test 1 vs Test 2;  $F(1,1933) = 354.63, p < 0.001$ ], speech training [Test 2 vs Test 3;  $F(1,1909) = 232.37, p < 0.001$ ], palate removal (Test 3 vs Test 4,  $F(1,1933) = 356.47, p < 0.001$ ] and the learning after-effect [Test 1 vs Test 4;  $F(1,1957) = 400.98, p < 0.001$ ]. All other principal effects, two-way interaction and three-way interaction effects were non-significant ( $p > 0.05$ ). The same pattern was observed for the tongue *blade* sensor, which showed a statistically significant fit for the effect of speech test associated with palate insertion [ $F(1,1933) = 346.87, p < 0.001$ ], speech training [ $F(1,1909) = 281.97, p < 0.001$ ], palate removal [ $F(1,1933) = 201.09, p < 0.001$ ] and the learning after-effect [ $F(1,1957) = 349.43, p < 0.001$ ]. All other principal effects, two-way interactions and three-way interaction effects were non-significant ( $p > 0.05$ ).

In summary, robust changes in tongue blade and dorsum position were observed in association with all four experimental effects, across all speech sounds and in all contexts.

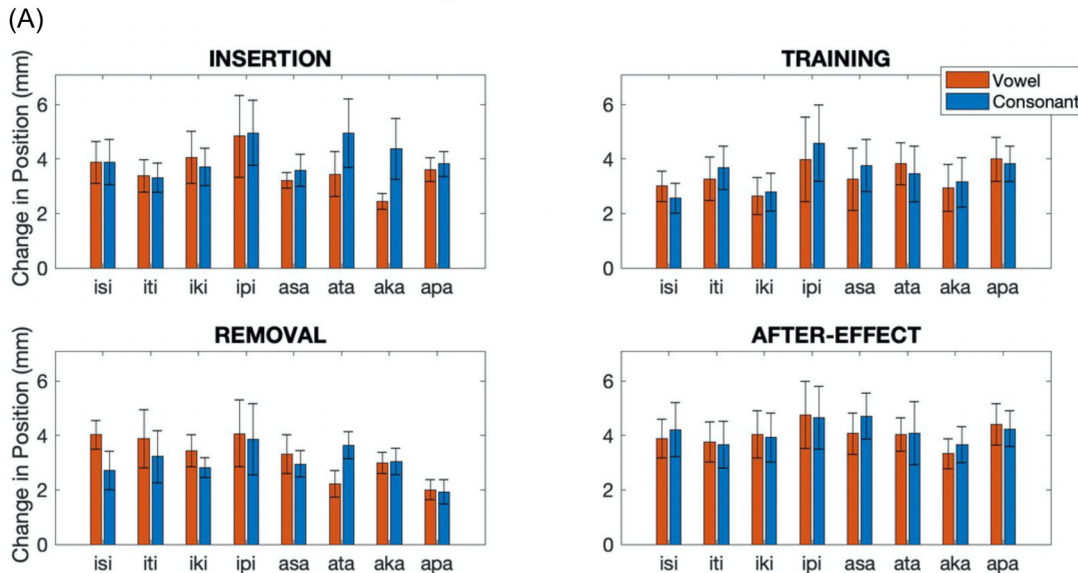
## D. Kinematics—Directional analyses

### 1. Effect of palate insertion and removal

The analyses of acoustic changes associated with palate insertion and removal revealed opposing effects across the two effects in both centroid frequency for consonants and formant frequency (primarily  $F2$ ) for vowels. The analysis of the kinematic amplitudes of tongue positional changes associated with palate insertion and removal further revealed the presence of reliable changes among all consonants and vowels. To further understand the nature of these kinematic changes, the *directions* of these tongue positional effects were examined. Figure 7(A) shows the mean direction (vector angle) of changes in tongue blade and dorsum position associated with palate insertion (small solid vectors) and removal (small dotted vectors), for each of the two vowels and four consonants (averaged across contexts). The grand mean direction of each effect is shown as a bold vector (insertion in blue, removal in red). Angles represent directions in the mid-sagittal plane, with  $0^\circ$  corresponding to the anterior direction (i.e., toward the subject's incisors) along the occlusal plane. The insertion effect shows an overall lowering of the tongue blade and advancement of the tongue dorsum sensor, consistent with a clockwise rotation of the tongue surface away from the palatal prosthesis. Immediately upon removal of the palate, the opposite pattern is observed on average: raising of the tongue blade and retraction of the dorsum sensor.

As can be seen in Fig. 7(A), the direction of tongue positional change associated with palate insertion and palate removal are nearly opposite to each other. Indeed, for the tongue blade sensor (left panel), the mean angular *difference* between the insertion and removal effects ranged from 135.7 to 173.2 degrees across the different speech sounds, with the grand average vectors differing by 157.9 degrees. For the tongue dorsum sensor, the angle differences ranged from 136.3 to 163.4 degrees across sounds, with the grand

### Tongue Blade Sensor



### Tongue Dorsum Sensor

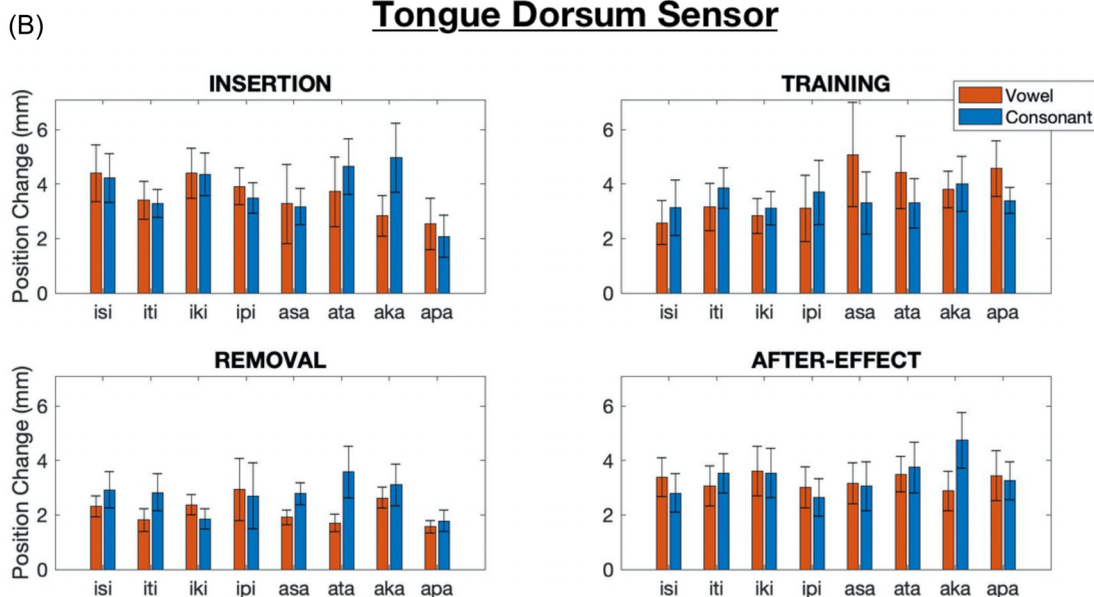


FIG. 6. (Color online) The mean change in position, associated with the four experimental effects of interest, of the tongue blade (top) and dorsum (bottom) sensors for the production of the syllable-initial vowel (red) and consonant (blue) among the eight different VCV contexts. Error bars show  $\pm 1$  standard error of the mean.

average vectors differing by 175.7 degrees. The reliability of differences in vector angles among the six speech sounds was examined using circular one-way ANOVA (*Watson-Williams* test), carried out separately within each of the two experimental effects. In both tongue sensors, no reliable difference was observed between speech sounds for the effect of palate insertion [*Blade*:  $F(5,48) = 0.97$ ,  $p = 0.45$ ; *Dorsum*:  $F(5,48) = 0.86$ ,  $p = 0.51$ ] or palate removal [*Blade*:  $F(5,48) = 0.05$ ,  $p = 0.99$ ; *Dorsum*:  $F(5,48) = 0.65$ ,  $p = 0.66$ ].

*Watson-Williams* tests were also used to assess the reliability of the difference in vector angle between the insertion and removal effects separately for each of the six sounds. In

the tongue blade sensor, with one exception, a statistically significant difference in the direction of tongue positional change between palate insertion and removal was observed for all sounds, including the two vowels [ $F(1,16) = 39.33$  [*i*]; 23.53 [*ae*],  $p < 0.05$  for all] and three consonants [ $F(1,16) = 10.63$  [*s*]; 68.18 [*k*]; 20.54 [*p*],  $p < 0.05$  for all]. The difference for the consonant /*t*/, however, was not significant [ $F(1,15) = 2.99$ ,  $p = 0.10$ ]. For the tongue dorsum sensor, the difference was significant for all sounds [ $F(1,16) = 38.61$  (*i*); 8.07 [*ae*]; 8.42 [*s*]; 9.85 [*t*]; 13.96 [*k*]; 10.53 [*p*],  $p < 0.05$  for all].

While averaging tongue positional changes across participants is useful to characterize overall trends, considerable

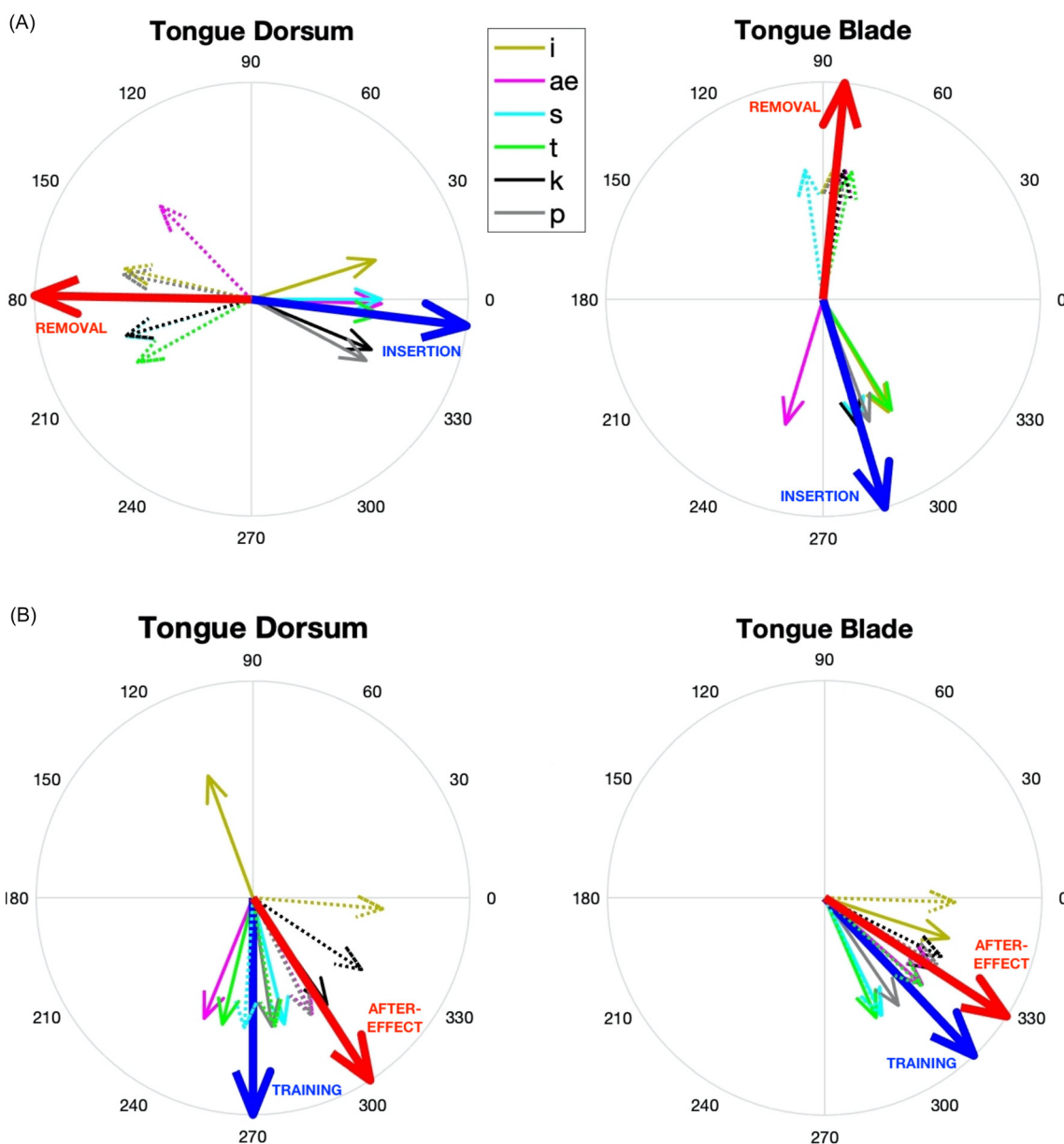


FIG. 7. (Color online) (A) Mean direction of tongue position change associated with palate insertion (small solid vectors) and removal (small dotted vectors), for each sound. The grand mean angle for each effect is shown as a large bold vector (insertion in blue, removal in red). Angles represent directions in the mid-sagittal plane, with  $0^\circ$  corresponding to the anterior direction (i.e., toward the subject's incisors) along the occlusal plane. The insertion effect shows an overall lowering of the tongue blade and advancement of the tongue dorsum. Immediately upon removal of the palate, the opposite pattern is observed on average. (B) Mean direction of tongue position change associated with speech training (small solid vectors) and the learning after-effect (small dotted vectors), for each sound. The grand mean angle for each effect is shown as a large bold vector (training in blue, after-effect in red). Both effects show a similar directional change in tongue position (for both sensors), characterized by an overall lowering and advancement of the tongue.

inter-speaker variation in oral kinematic patterns is a common finding in studies of speech production, and has been specifically noted in prior studies involving adaptation to palatal prostheses (e.g., Hamlet *et al.*, 1978; Baum and McFarland, 2000; Thibeault *et al.*, 2011). In the current study, therefore, a key question remains whether different participants may have responded to the palatal perturbation with varying directional changes in tongue kinematics. In the case of circular (angular) data, one approach to addressing such questions is through a test of directional non-uniformity (e.g., the *Raleigh z* test) testing the hypothesis that a set of vectors is not distributed

evenly around a circle. For the effect of palate insertion, a test of non-uniformity among the participants was carried out separately for each target sound, including the two word-initial vowels (averaged across consonant contexts) and the four consonants (averaged across vowel contexts). For the tongue blade sensor, the *Rayleigh* tests yielded a non-significant result (i.e., not rejecting the null hypothesis of directional uniformity) for the consonant /s/ ( $z = 1.70$ ,  $p = 0.18$ ). In contrast, a significant departure from uniformity was found for the consonants /t/ ( $z = 3.56$ ,  $p < 0.05$ ), /k/ ( $z = 8.45$ ,  $p < 0.05$ ) and /p/ ( $z = 5.75$ ,  $p < 0.05$ ), as well as the

vowels /i/ ( $z=4.05$ ,  $p<0.05$ ) and /æ/ ( $z=4.16$ ,  $p<0.05$ ). For the tongue dorsum sensor, a similar pattern was observed with a non-significant result for /s/ ( $z=0.69$ ,  $p=0.52$ ) and significant departures from uniformity for /t/ ( $z=2.48$ ,  $p<0.05$ ), /k/ ( $z=6.73$ ,  $p<0.05$ ), /p/ ( $z=4.01$ ,  $p<0.05$ ), /i/ ( $z=8.29$ ,  $p<0.05$ ) and /æ/ ( $z=1.42$ ,  $p<0.05$ ). These results suggest that, following insertion of the palatal prosthesis, sounds that strongly interact with the palate (e.g., /s/) may show a wider range of directional patterns among participants, whereas for the remaining consonants and vowels, the directions of tongue kinematic changes may be more systematic among talkers.

## 2. Effects of speech training and the learning after-effect

The analysis of kinematic amplitude indicated robust changes in tongue position for the production of all consonants and vowels following 20 min of speech training, as well as significant after-effects that persisted following removal of the palatal prosthesis. As with the effects of palate insertion and removal, understanding the nature of these training-related changes requires a further examination of their directional patterns. Figure 7(B) shows the mean direction of tongue positional change associated with speech training and learning after-effects for each vowel and consonant, along with the grand average across all sounds. The effect of training is generally characterized by an advancement and lowering of the tongue, with the after-effect showing a similar directional pattern.

The reliability of differences in vector angles across the six speech sounds was examined using circular one-way ANOVA. For the tongue blade sensor, no significant difference was observed across speech sounds for the effect of speech training [ $F(5,48)=0.22$ ,  $p=0.95$ ] or for the after-effect [ $F(5,48)=0.36$ ,  $p=0.87$ ]. Similarly for the tongue dorsum sensor, no significant difference was observed across speech sounds for the effect of training [ $F(5,48)=0.83$ ,  $p=0.54$ ] or for the after-effect [ $F(5,48)=1.01$ ,  $p=0.42$ ].

If changes in the positioning of the tongue observed following speech training are indeed a reflection of changes in feed-forward speech motor planning (i.e., motor learning), the directions of tongue kinematic changes associated with the *after-effect* (following removal of the perturbation) would be predicted to be consistent with those training-related effects. The average patterns shown in Fig. 7(B) are broadly consistent with this prediction (i.e., for each sound considered individually, as well as for the grand average directions for each sensor, the angular difference between training and after-effects are near zero). For the tongue blade sensor, the angular difference between training and after-effects was close to zero for all speech sounds, ranging from 0.4 to 25.1 degrees. The angle difference between grand average vectors for the training and after-effect was similarly small, at just 11.3 degrees. For the tongue dorsum sensor, with a single exception (the vowel /i/), the change in position associated with the training and after-effects were similarly directionally aligned, with angular differences

ranging from 17.8 to 49.0 degrees across the different speech sounds (the sound /i/ showed a somewhat larger difference of 115.1 degrees), and the grand average vectors differing by 35.4 degrees.

*Watson-Williams* tests were used to assess the reliability of the difference in vector angles between the training and after-effects, separately for each of the six sounds. In the tongue blade sensor, the difference in the direction of tongue positional change was not statistically significant for all sounds [ $F(1,16)=0.14$  [i]; 0.02 [ae]; 0.01 [s]; 0.08 [t]; 0.03 [k]; 0.37 [p],  $p>0.05$  for all]. The tongue dorsum sensor similarly showed no significant differences in the direction of tongue position change between the training and after-effects in the case of all sounds [ $F(1,16)=2.03$  [i]; 0.56 [ae]; 0.10 [s]; 0.18 [t]; 0.12 [k]; 0.13 [p],  $p>0.05$  for all].

The possibility of talker-specific differences in tongue kinematic patterns remains highly relevant in the context of the effects related to speech training (and their corresponding after-effects), where different participants might reasonably be expected to use a range of articulatory strategies to compensate for the individually varying effects of the palatal perturbation. Figure 8 provides an illustration of such variability in the current study, showing the directions of tongue position change (tongue blade sensor) for all participants (different colored vectors) associated with the four consonants in each vowel context. Widely varying directional patterns among the talkers are observed in each case. *Rayleigh z* tests were used once again to examine the variation in directional changes among the participants during production of each of the two word-initial vowels and each of the four consonants. For the tongue blade sensor, no significant departure from uniformity was found for any sound [[i]:  $z=0.18$ ,  $p=0.84$ ; [ae]:  $z=0.03$ ,  $p=0.97$ ; [s]:  $z=1.72$ ,  $p=0.18$ ; [t]:  $z=1.21$ ,  $p=0.31$ ; [k]:  $z=1$ ,  $p=0.38$ ; [p]:  $z=0.16$ ,  $p=0.86$ ]. The same pattern was observed for the tongue dorsum sensor [[i]:  $z=0.25$ ,  $p=0.79$ ; [ae]:  $z=0.51$ ,  $p=0.62$ ; [s]:  $z=0.13$ ,  $p=0.88$ ; [t]:  $z=0.69$ ,  $p=0.52$ ; [k]:  $z=0.06$ ,  $p=0.94$ ; [p]:  $z=0.16$ ,  $p=0.86$ ], confirming the highly variable nature of the articulatory strategies associated with training among the different participants.

The above analyses reveal two important effects of speech training on tongue kinematic patterns: (1) the overall similarity in direction of tongue positional change associated with the training and after-effects [illustrated by the *convergent* mean vector directions between the two effects shown in Fig. 7(B)], and (2) the considerable variability among individual talkers in the tongue kinematic strategies used to adapt to the perturbation (illustrated by the *divergent* directions of tongue positional change among participants shown in Fig. 8). If the observed changes in tongue position following speech training genuinely reflect learned changes in articulatory planning (i.e., changes in feed-forward control), they should be related to the directions of the observed after-effects regardless of inter-talker differences in the precise patterning of those training effects. To examine the global directional relationship between the training and after-effects, the *angular difference* between the training

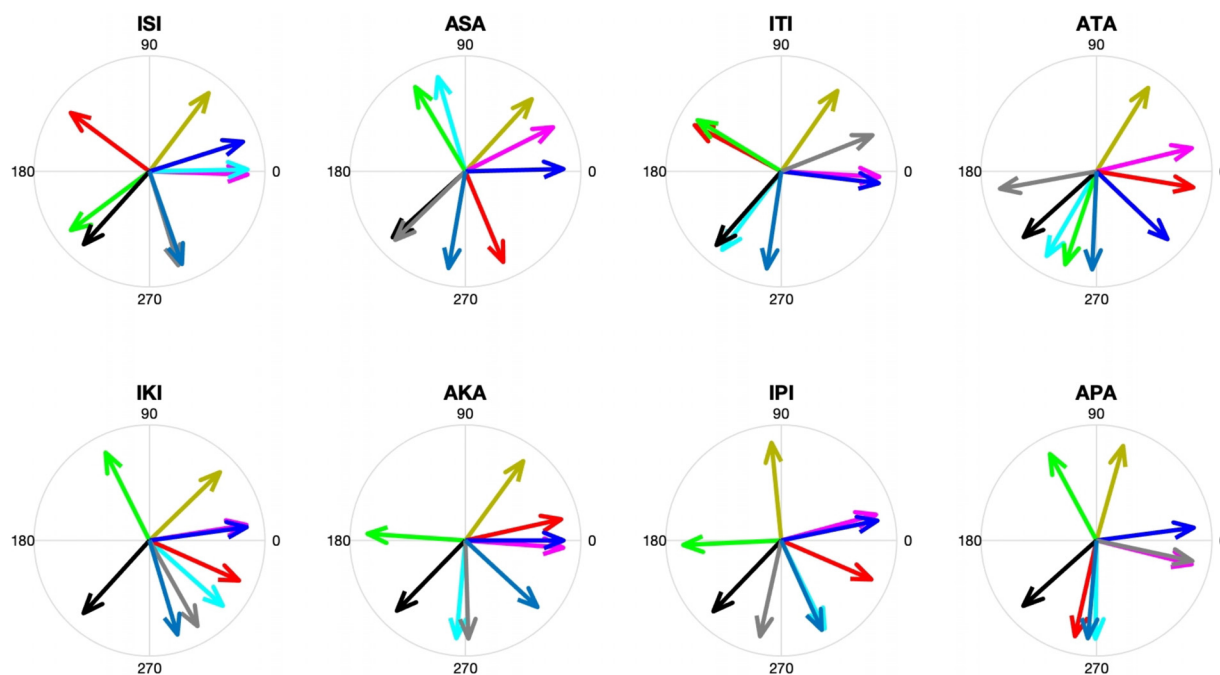


FIG. 8. (Color online) Directions of changes in tongue blade position for each participant (different colored vectors) associated with production of the consonant in each of the eight VCV contexts. A wide range of directions can be observed among the participants.

and after-effects was calculated separately for each speech sound, within each VCV context, and within each of the participants, and then pooled together to consider the overall trend among the group. Figure 9 shows polar histograms of the distributions of angle differences related to the production of vowels (left panel) and consonants (right panel) across all VCV contexts and participants. The plots reveal that even when all participants and contexts are considered individually, the training effects and learning after-effects

remain characterized by an average directional difference close to 0° for both the consonants (11.7° for tongue blade and 14.9° for tongue dorsum) and vowels (1.88° for tongue blade and -7.8° for tongue dorsum). Further, *Rayleigh z* tests of non-uniformity on the sets of angular differences were highly significant for the tongue blade sensor (vowels:  $z = 41.4, p < 0.001$ ; consonants:  $z = 36.9, p < 0.001$ ) and the tongue dorsum (vowels:  $z = 20.2, p < 0.001$ ; consonants:  $z = 19.67, p < 0.001$ ).

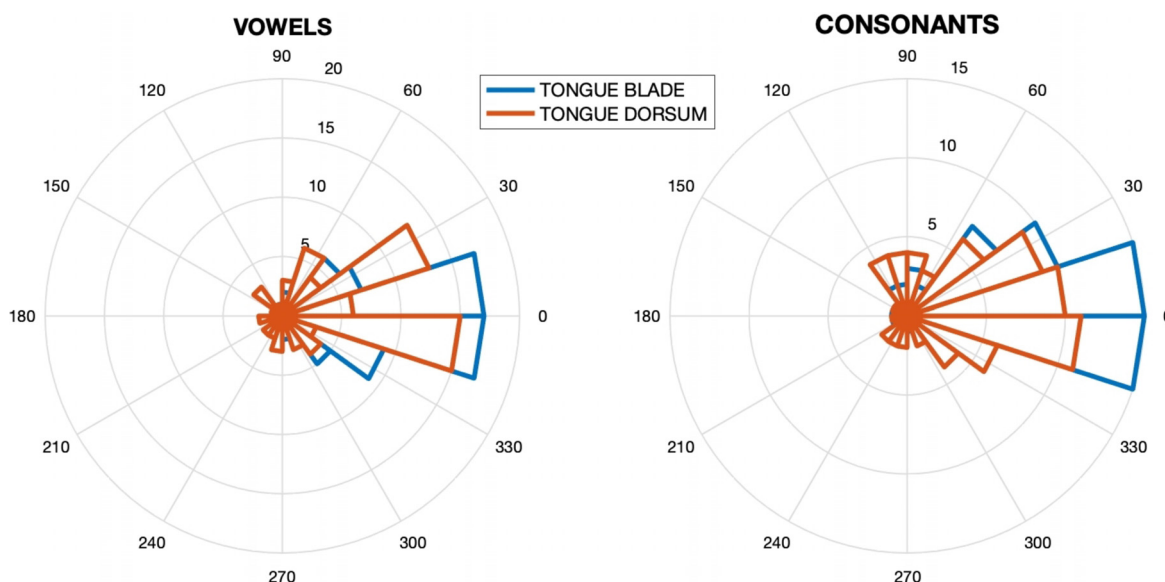


FIG. 9. (Color online) Polar histograms showing the distributions of angle differences between speech training and after-effects, related to the production of vowels (left panel) and consonants (right panel) among all participants and VCV contexts. Data are shown for both the tongue blade (blue) and dorsum (orange) sensors. Even when participants and contexts are considered on an individual basis, the training effects and learning after-effects remain characterized by an average directional difference close to 0°.

#### IV. DISCUSSION

The present study examined the adaptation of speech production to a prosthesis that altered the shape of the hard palate during the production of a range of consonants and vowels. Acoustically, as demonstrated in a number of prior studies, a reliable perturbation and training effect was observed for the production of sounds directly implicating the altered alveolar region (/s/ and /t/). In addition, an acoustic effect of the perturbation (immediately following insertion and removal of the palatal prosthesis) was observed during the production of the velar consonant /k/, as well as the vowels /i/ and /æ/, with the production of the vowels also showing a small but statistically reliable effect of speech training in  $F2$  frequency. The pattern of results illustrates the non-uniform impact of the palatal perturbation on speech acoustic properties, while at the same time suggesting that the effects related to the perturbation and subsequent training extend beyond those sounds directly involving physical interaction with the palate (specifically, the velar stop consonant /k/ and low vowel /æ/). At the level of tongue kinematics, the effects of the palatal perturbation were found to be more pronounced, with a reliable change in tongue position observed across *all* the tested sounds following palate insertion and removal. Subsequent speech training—involving the production of complex, variable sentences—was also associated with robust changes in tongue position across all sounds and contexts. Strikingly, a comparable amplitude of tongue positional change was observed among the various speech sounds, despite differences in the degree of physical interaction with the palate.

At the level of kinematics, the analyses revealed the global nature of the effects of the palatal perturbation, with a reliable change in tongue position observed across all of the tested sounds following palate insertion and removal. Insertion of the palate induces immediate changes in tongue position characterized by an overall lowering of the blade and advancement of the dorsum sensor, with the opposite pattern observed immediately upon removal of the palate. Following 20 min of sentence production with the palate in place, an overall advancement and lowering of the tongue is observed across the different sounds, with a similar directional pattern observed for the after-effect. While these average tongue movement patterns are helpful for characterizing the kinematic effects in broad terms, the directional patterns associated with the perturbation, as well as the subsequent adaptation and after-effects, were found to be, not surprisingly, extremely variable among participants. Critically, despite this directional variability, the vector angles associated with speech training among the varying sounds, contexts, and participants were found to be individually consistent with the corresponding directions of each learning after-effect, providing a strong indication that those kinematic changes were the result of adaptive changes in feed-forward control (i.e., motor learning). The acoustic and kinematic findings in the present study therefore support the idea that talkers are able to simultaneously adapt multiple

speech sounds (including vowels, fricatives, and stop consonants) across the articulatory workspace to a physical perturbation impacting both somatosensory and auditory feedback. The results thus provide an important validation and extension of the result of Lametti *et al.* (2018), in which speech adaptation across the workspace was observed for vowels in response to a uniform, auditory-only spectral shift.

In the present study, acoustic changes associated with insertion of the palate were found to be non-uniform across the workspace, with differing degrees of spectral changes observed among the different consonant and vowel sounds. The perturbation, however, induced a comparable magnitude of *kinematic* change in speech motor patterns across all sounds. This discrepancy reflects, in part, the complex non-linear relationship between physical vocal tract configurations and acoustic outcomes. Following 20 min of speech practice, small but reliable changes in speech acoustics were observed related to the production of only a subset of sounds (in particular, the fricative /s/). Robust adaptive changes in tongue movement patterns measured kinematically, however, were observed once again across *all* speech sounds. The finding of motor adaptation across the articulatory workspace despite differences in acoustic consequences among the different speech sounds suggests that the adaptation of speech motor patterns was not driven exclusively by auditory feedback, but rather was related, at least in part, to perceived changes in somatosensory feedback. The role of somatosensory feedback has been previously highlighted in studies of speech adaptation (Tremblay *et al.*, 2003; Nasir and Ostry, 2006; Lametti *et al.*, 2014); however, such studies have typically focused on a limited number of speech sound targets within narrow phonetic contexts at any one time. Here, robust speech adaptation was revealed across multiple sounds following training that involved the production of complex, variable sentences. As such, the current results support a role for somatosensory feedback in speech motor learning under conditions that more closely resemble naturalistic, real-world speaking conditions.

A prior study by Honda *et al.* (2002; see also Honda and Kaburagi, 2000) also highlighted the role of somatosensory feedback in compensatory changes to a change in palatal shape, using an inflatable device that dynamically altered the thickness of an artificial palate during the repeated production of CV syllables /ʃa/ or /tʃa/ in two adult talkers. In addition to acoustic and perceptual measures, mid-sagittal tongue position was tracked using electromagnetic articulography (EMA). The perturbation was unexpectedly applied immediately prior to syllable onset, and then maintained for several additional repetitions of the syllable, in order to assess immediate feedback-related effects, as well as an effect of brief practice. Compensatory changes were observed kinematically within the first syllable; however, they were incomplete (resulting in positional overshoot), leading to acoustically and perceptually measurable distortions during consonant production. Compensation continued through subsequent syllable productions, improving the quality of the acoustic output. Critically, significant

immediate and subsequent compensatory changes were observed even in the presence of auditory masking, indicating a key role for somatosensory feedback in the compensatory response.

The kinematic analyses in the present study provide potentially novel insights into the effects of a local structural perturbation of the vocal tract on the control of oral articulatory movements during speech production. Past studies examining palatal perturbations have typically focused on sounds that interact directly with the palate (e.g., /s/, /t/), reflecting the idea that such a perturbation was likely to be local in nature (e.g., Hamlet *et al.*, 1978; Baum and McFarland, 1997, 2000; Aasland *et al.*, 2006). This view is certainly reasonable, given the results earlier studies noting only limited acoustic perturbation and adaptation effects in sounds that did not involve direct contact with the structural modification (e.g., McFarland *et al.*, 1996). The present kinematic results, however, paint a different picture of the perturbing effects of a local oral anatomical change on speech production. Immediately following insertion of the palatal prosthesis, changes in tongue position were observed in the production of *all* speech sounds, characterized by a global tongue lowering and retracting movement (though, with considerable variability in kinematic patterns among participants). Given the lack of physical contact between the tongue and palate during the production of many of these sounds (/k/, /p/, /æ/), such relatively sudden changes appear to reflect an *active* motor process that impacts tongue motor patterns during the production of all sounds in response to the altered palatal shape, rather than a simple physical interaction. It is interesting to consider how such a global mechanism of motor adjustment relates to the more gradual and fine-tuned process of adaptation that occurs with speech practice under altered sensory conditions. Future studies, combining different physical manipulations, kinematic measures of the oral articulators and neuroimaging, may help to elucidate the nature of the sensory and motor neural processes that underlie both types of speech motor adjustment.

While the majority of prior studies examining adaptation to a palatal perturbation have focused on a narrow set of sounds and contexts using acoustic measures, one study by Brunner (2009; also published in Brunner *et al.*, 2007, 2011) notably examined a wider range of phonemes including fricatives, stop consonants and vowels in seven talkers who adapted their speech to a palatal prosthesis over a period of two weeks. Speech patterns were examined both acoustically and kinematically (using EMA) immediately following the initial insertion of the palate (with and without masking noise), and then after one week and two weeks of practice. The interpretation of the resulting changes was complicated somewhat by the use of two different shapes of prostheses among the participants, with three subjects wearing one that flattened the central arch, and four subjects wearing one that lowered the alveolar ridge. Acoustic and kinematic data were analyzed separately for each subject, providing a nuanced but complex picture of the resulting perturbation and training effects. In general, for vowels,

both acoustic (formants) and kinematic measures were found to be influenced by the presence of the palate as well as the period of speech practice, however the practice-related changes were reported to be highly variable among the subjects and did not systematically involve a return to baseline values. Fricatives involving direct interactions with the palate (notably, the alveolar fricative) appeared to show more systematic acoustic and kinematic changes in response to the perturbation and subsequent practice; however, no systematic characterization of the kinematic changes for the remaining fricatives or the stop consonants emerged. While these observations of Brunner (2009) are broadly compatible with the observed patterns in the current study, the ability to precisely measure the changes in tongue kinematics across key phases (e.g., to compare tongue positions before and after practice with the palate in place, or to compare pre-insertion baseline with post-training after-effect) was not possible due to the fact that the EMA sensors had to be removed and re-attached between recording sessions. While this did not prevent the authors from exploring a number of hypotheses concerning the control of acoustic/perceptual contrast, higher derivatives of motion (e.g., tangential jerk), and trading-relations in kinematic patterns, the question of whether the practice-related kinematic changes reflected learned changes in feedforward control across the range of different speech sounds was never directly addressed.

An appeal of studying sensorimotor adaptation to a *physical* manipulation of the vocal tract (compared to a focused manipulation of auditory feedback, for example) is its potentially closer link with real-world experiences. There is evidence, for example, that talkers adjust their tongue kinematics to accommodate naturally occurring differences in palatal morphology. Lammert *et al.* (2014) examined the production of the high-front vowel /i/ in five talkers using real-time MRI, with the aim of relating naturally-occurring differences in palate shape with variation in acoustic (formants) and mid-sagittal tongue kinematic patterns. They found that talkers indeed adapted their tongue shape during vowel production to accommodate aspects of palatal morphology (anteriority and concavity) that were predicted to have a significant impact on vowel formants. Other studies have similarly shown that for speech sounds typically produced with an elevated tongue body (such as high vowels, or the consonant /s/), the positioning of the tongue shows sensitivity to differences in the height of the palatal arch (Grimm *et al.*, 2017; Hasegawa Johnson *et al.*, 2003; Weirich and Fuchs, 2013), as well as the slope of the anterior palate (Stone *et al.*, 2019). It has also been demonstrated that talkers with naturally flatter palates exhibit reduced tongue kinematic variability, possibly reflecting the increased sensitivity of vowel acoustics to small changes in tongue position associated with that palatal shape (Brunner *et al.*, 2009, Mooshammer *et al.*, 2004).

Parallels can also be made to the complex non-linear growth of the vocal apparatus during childhood (Vorperian *et al.*, 2009; Barbier *et al.*, 2015), or other oral structural changes resulting from tooth loss, orthodontic implants or

dentures, or maxillofacial surgery. Early clinical studies suggested that talkers require weeks or months to adapt to structural modifications of the oral cavity, such as dental implants (Troffer and Beder, 1961; Bergman and Carlsson, 1972). Indeed, in the earliest experimental investigations of adaptation to an artificial palate, talkers were exposed to the manipulation for 1–2 weeks in order to examine compensatory strategies over this expected period of adaptation (Hamlet *et al.*, 1976; Hamlet *et al.*, 1978). Brunner *et al.* (2009) also examined adaptation over a period of two weeks. Other recent studies of adaptation to palatal perturbations have shown reliable adaptation effects within the much shorter time-frame of 15–20 min (McFarland *et al.*, 1996; Baum and McFarland, 1997). However, as noted earlier, the observed training effects in these studies tended to be limited in scope, with the effects restricted to sounds involving direct interactions with the palate (e.g., /s/). The kinematic measures in the present study extend these earlier findings of speech adaptation within narrow contexts by demonstrating that talkers are in fact capable of robust speech motor adaptation across a wide range of sounds and contexts following just 20 min of practice. Given the physical nature of the oral perturbation and the complex nature of the practice task (highly variable, complex sentences), the current results strongly suggest that talkers are likely to be capable of speech motor adaptation within a similar time-frame to a wide range of “real-world” structural modifications, such as dental implants. Future studies using kinematic measures to study speech adaptation in such clinical contexts would be very helpful in gaining deeper insights into such effects.

## ACKNOWLEDGMENTS

This study was supported by grants from the Natural Sciences and Engineering Council of Canada (NSERC-Canada) and the Centre for Research on Brain, Language and Music (CRBLM). We also thank Sabine Burfin for the illustration in Fig. 1.

<sup>1</sup>It is important to distinguish studies that have examined motor responses to sudden, individual, unexpected sensory perturbations versus those that have employed predictable sensory-feedback alterations that are maintained for a period of time (minutes, hours or even days). The former allows for the study of rapid online, feedback-based responses to perceived sensory errors within a given speech utterance (often referred to as “compensation”; e.g., Burnett *et al.*, 1998; Purcell and Munhall, 2006), whereas the latter allows for the study of more gradually induced changes in the predictive, feed-forward control of speech movements resulting from a more substantial period of practice (often using the term “sensorimotor adaptation” or “speech motor learning”). The present study is interested specifically in processes of sensorimotor adaptation.

<sup>2</sup>Open Science Framework repository available at <https://osf.io/>.

<sup>3</sup>See supplementary material at <https://doi.org/10.1121/10.0000672> for the complete text output of the linear mixed effects regression models presented in this paper.

Aasland, W. A., Baum, S. R., and McFarland, D. H. (2006). “Electropalatographic, acoustic, and perceptual data on adaptation to a palatal perturbation,” *J. Acoust. Soc. Am.* **119**(4), 2372–2381.  
Barbier, G., Boë, L. J., Captier, G., and Laboissière, R. (2015). “Human vocal tract growth: A longitudinal study of the development of various

anatomical structures,” in *Proceedings of the Sixteenth Annual Conference of the International Speech Communication Association (INTERSPEECH 2015)*, pp. 364–368.

- Baum, S. R., and McFarland, D. H. (1997). “The development of speech adaptation to an artificial palate,” *J. Acoust. Soc. Am.* **102**(4), 2353–2359.  
Baum, S. R., and McFarland, D. H. (2000). “Individual differences in speech adaptation to an artificial palate,” *J. Acoust. Soc. Am.* **107**(6), 3572–3575.  
Bates, D., Maechler, M., Bolker, B., and Walker, S. (2019). “lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-21,” <http://CRAN.R-project.org/package=lme4> (downloaded November 1, 2019).  
Berens, P. (2009). “CircStat: A MATLAB toolbox for circular statistics,” *J. Stat. Softw.* **31**(10), 1–21.  
Bergman, B., and Carlsson, G. E. (1972). “Review of 54 complete denture wearers: Patients opinions 1 year after treatment,” *Acta. Odontol. Scand.* **30**(4), 399–414.  
Brunner, J. (2009). “Perturbed Speech: How compensation mechanisms can inform us about 1154 phonemic targets,” *Sudwestdeutscher Verlag Fur Hochschulschriften* **196**, hal-00372151.  
Brunner, J., Fuchs, S., and Perrier, P. (2009). “On the relationship between palate shape and articulatory behavior,” *J. Acoust. Soc. Am.* **125**(6), 3936–3949.  
Brunner, J., Hoole, P., and Perrier, P. (2007). “Articulatory optimisation in perturbed vowel articulation,” in *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS 2007)*, August 6–10, Saarbrücken, Germany, pp. 497–500.  
Brunner, J., Hoole, P., and Perrier, P. (2011). “Adaptation strategies in perturbed /s/,” *Clin. Ling. Phon.* **25**, 705–724.  
Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (1998). “Voice F0 responses to manipulations in pitch feedback,” *J. Acoust. Soc. Am.* **103**(6), 3153–3161.  
Cai, S., Ghosh, S. S., Guenther, F. H., and Perkell, J. S. (2011). “Focal manipulations of formant trajectories reveal a role of auditory feedback in the on-line control of both within-syllable and between-syllable speech timing,” *J. Neurosci.* **31**, 16483–16490.  
Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. N. (1988). “Statistical analysis of word-initial voiceless obstruents: Preliminary data,” *J. Acoust. Soc. Am.* **84**(1), 115–123.  
Garber, S. R., Speidel, T. M., Siegel, G. M., Miller, E., and Glass, L. (1980). “The effects of presentation of noise and dental appliances on speech,” *J. Speech Hear. Res.* **23**, 838–852.  
Gay, T., Lindblom, B., and Lubker, J. (1981). “Production of bite-block vowels: Acoustic equivalence by selective compensation,” *J. Acoust. Soc. Am.* **69**(3), 802–810.  
Grimm, D. L., Stone, M., Woo, J., Lee, J., Hwang, J. H., Bedrosian, G. E., and Prince, J. L. (2017). “The effects of palate features and glossectomy surgery on /s/ production,” *J. Speech Lang. Hear. Res.* **60**(12), 3417–3425.  
Hamlet, S., Geoffrey, V., and Bartlett, D. (1976). “Effect of the dental prosthesis on speaker specific characteristics of voice,” *J. Speech Hear. Res.* **19**(4), 639–650.  
Hamlet, S., Stone, M., and McCarty, T. (1978). “Conditioning prostheses viewed from the standpoint of speech adaptation,” *J. Prosthet. Dent.* **40**(1), 60–66.  
Hasegawa Johnson, M., Pizza, S., Alwan, A., Cha, J. S., and Haker, K. (2003). “Vowel category dependence of the relationship between palate height, tongue height, and oral area,” *J. Speech Lang. Hear. Res.* **46**, 738–753.  
Honda, M., Fujino, A., and Kaburagi, T. (2002). “Compensatory responses of articulators to unexpected perturbation of the palate shape,” *J. Phon.* **30**(3), 281–302.  
Honda, M., and Kaburagi, T. (2000). “Speech compensation to dynamical structural perturbation of the palate shape,” in *Proceedings of the 5th Speech Production Seminar and CREST Workshop on Models of Speech Production*, May 1–4, Kloster Seon, Germany, pp. 21–24.  
Houde, J. F., and Jordan, M. I. (1998). “Sensorimotor adaptation in speech production,” *Science* **279**(5354), 1213–1216.  
Houde, J. F., and Nagarajan, S. S. (2011). “Speech production as state feedback control,” *Front. Hum. Neurosci.* **5**(82).  
IEEE (1969). “Recommended practice for speech quality measurements,” *IEEE Trans. Audio Electroacoust.* **17**, 227–246.



- Johnson, A. A., Reidy, P. F., and Edwards, J. R. (2018). "Quantifying robustness of the /t-/k/ contrast using a single, static spectral feature," *J. Acoust. Soc. Am.* **144**(2), 105–111.
- Jones, J. A., and Munhall, K. G. (2003). "Learning to produce speech with an altered vocal tract: The role of auditory feedback," *J. Acoust. Soc. Am.* **113**(1), 532–543.
- Kelso, J. A. S., Tuller, B., Bateson, E. V., and Fowler, C. A. (1984). "Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures," *J. Exp. Psychol. Hum. Percept Perform.* **10**, 812–832.
- Lametti, D. R., Nasir, S. M., and Ostry, D. J. (2012). "Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback," *J. Neurosci.* **32**(27), 9351–9358.
- Lametti, D. R., Rochet-Capellan, A., Neufeld, E., Shiller, D. M., and Ostry, D. J. (2014). "Plasticity in the human speech motor system drives changes in speech perception," *J. Neurosci.* **34**(31), 10339–10346.
- Lametti, D. R., Smith, H. J., Watkins, K. E., and Shiller, D. M. (2018). "Robust sensorimotor learning during variable sentence-level speech," *Curr. Biol.* **28**(19), 3106–3113.
- Maniwa, K., Jongman, A., and Wade, T. (2009). "Acoustic characteristics of clearly spoken English fricatives," *J. Acoust. Soc. Am.* **125**(6), 3962–3973.
- McFarland, D. H., Baum, S. R., and Chabot, C. (1996). "Speech compensation to structural modifications of the oral cavity," *J. Acoust. Soc. Am.* **100**, 1093–1104.
- Mooshammer, C., Perrier, P., Fuchs, S., Geng, C., and Pape, D. (2004). "An EMMA and EPG study on token-to-token variability," *AIPUK* **36**, 47–63.
- Nasir, S., and Ostry, D. (2006). "Somatosensory precision in speech production," *Curr. Biol.* **16**(19), 1918–1923.
- Parrell, B., Lammert, A. C., Ciccarelli, G., and Quatieri, T. F. (2019). "Current models of speech motor control: A control-theoretic overview of architectures and properties," *J. Acoust. Soc. Am.* **145**(3), 1456–1481.
- Parrell, B., Ramanarayanan, V., Nagarajan, S., and Houde, J. (2019). "The FACTS model of speech motor control: Fusing state estimation and task-based control," *PLoS Comput. Biol.* **15**(9), e1007321.
- Perkell, J. S. (2012). "Movement goals and feedback and feedforward control mechanisms in speech production," *J. Neuroling.* **25**(5), 382–407.
- Purcell, D. W., and Munhall, K. G. (2006). "Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation," *J. Acoust. Soc. Am.* **120**(2), 966–977.
- Ramanarayanan, V., Parrell, B., Goldstein, L., Nagarajan, S., and Houde, J. (2016). "A new model of speech motor control based on task dynamics and state feedback," in *Proceedings of the Seventeenth Annual Conference of the International Speech Communication Association (INTERSPEECH 2016)*, September 8–12, San Francisco, CA, pp. 3564–3568.
- R Core Team (2019). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org/> (downloaded November 1, 2019).
- Savariaux, C., Perrier, P., and Orliaguet, J. P. (1995). "Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: A study of the control space in speech production," *J. Acoust. Soc. Am.* **98**(5), 2428–2442.
- Shiller, D. M., Sato, M., Gracco, V. L., and Baum, S. R. (2009). "Perceptual recalibration of speech sounds following speech motor learning," *J. Acoust. Soc. Am.* **125**(2), 1103–1113.
- Smith, A., and Zelaznik, H. N. (2004). "Development of functional synergies for speech motor coordination in childhood and adolescence," *Dev. Psychobiol.* **45**(1), 22–33.
- Stone, M., Gomez, A. D., Zhuo, J., Tchouaga, A. L., and Prince, J. L. (2019). "Quantifying tongue tip shape in apical and laminal /s/: Contributions of palate shape," *J. Speech Lang. Hear. Res.* **62**(9), 3149–3159.
- Thibeault, M., Ménard, L., Baum, S. R., Richard, G., and McFarland, D. H. (2011). "Articulatory and acoustic adaptation to palatal perturbation," *J. Acoust. Soc. Am.* **129**(4), 2112–2120.
- Tourville, J. A., and Guenther, F. H. (2011). "The DIVA model: A neural theory of speech acquisition and production," *Lang. Cogn. Process.* **26**, 952–981.
- Tremblay, S., Shiller, D. M., and Ostry, D. J. (2003). "Somatosensory basis of speech production," *Nature* **423**(6942), 866–869.
- Troffer, C., and Beder, O. E. (1961). "Immediate dentures and speech defects," *Dent. Prog.* **1**, 264.
- Villacorta, V. M., Perkell, J. S., and Guenther, F. H. (2007). "Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception," *J. Acoust. Soc. Am.* **122**(4), 2306–2319.
- Vorperian, H. K., Wang, S., Chung, M. K., Schimek, E. M., Durtschi, R. B., Kent, R. D., and Gentry, L. R. (2009). "Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study," *J. Acoust. Soc. Am.* **125**(3), 1666–1678.
- Weirich, M., and Fuchs, S. (2013). "Palatal morphology can influence speaker-specific realizations of phonemic contrasts," *J. Speech Lang. Hear. Res.* **56**, 1894–1908.