Speaking tongues are actively braced

Bryan Gick\textsuperscript{a}, Blake Allen\textsuperscript{a}, Francois Roewer-Despres\textsuperscript{b}, Ian Stavness\textsuperscript{*b}

\textsuperscript{a} Department of Linguistics, University of British Columbia, Vancouver, British Columbia, Canada

\textsuperscript{b} Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

* Corresponding author:
Ian Stavness
Department of Computer Science, University of Saskatchewan, 176 Thorvaldson Building, 110 Science Place, Saskatoon, SK S7N 5C9 Canada

Phone: 306-966-7995
Email: ian.stavness@usask.ca

Keywords
speech production, tongue bracing, electropalatography, biomechanics, articulation
Abstract

**Purpose:** Bracing of the tongue against opposing vocal tract surfaces such as the teeth or palate has long been discussed in the context of biomechanical, somatosensory and aeroacoustic aspects of tongue movement. However, previous studies have tended to describe bracing only in terms of contact (rather than mechanical support), and only in limited phonetic contexts, supporting a widespread view of bracing as an occasional state, peculiar to specific sounds or sound combinations.

**Methods:** The present study tests the pervasiveness and effortfulness of tongue bracing in continuous English speech passages using Electropalatography (EPG) and 3D biomechanical simulations.

**Results:** The tongue remains in continuous contact with the upper molars during speech, with only rare exceptions. Use of the term “bracing” (rather than merely “contact”) is supported here by biomechanical simulations showing that lateral bracing is an active posture requiring dedicated muscle activation; further, loss of lateral contact for onset /l/ allophones is found to be consistently accompanied by contact of the tongue blade against the anterior palate. In the rare instances where direct evidence for contact is lacking (only in a minority of low vowel and postvocalic /l/ tokens), additional biomechanical simulations show that contact is maintained against lateral pharyngeal structures dorsal to the teeth.

**Conclusions:** Taken together, these results indicate that tongue bracing is both pervasive and active in running speech and essential in understanding tongue movement control.
1. Introduction

The term “bracing” has often been used to describe a lingual posture in speech wherein the tongue is in contact with a rigid vocal tract surface, such as the teeth or palate. While “bracing” implies active mechanical support rather than mere contact, tongue bracing has typically been both defined and measured strictly in terms of contact. For example, McLeod et al. (2006: 52) describe “lateral contact along the length of the palate (termed lateral bracing)” while Gibbon et al. (2010: 406) observe that “the lateral margins of the tongue make contact against the palate, a phenomenon known as lateral bracing (Stone, 1991).” This usage is often applied in papers employing experimental tools such as electropalatography, which measures only contact, e.g., Narayanan et al. (1997: 1070) state: “EPG data…consistently showed linguo-alveolar bracing.” The central goal of the present paper is to seek evidence testing whether the lingual contact observed in these studies in fact constitutes “bracing” in the sense that it is intentional and effortful, and to determine whether this evidence corroborates the widespread view of bracing as providing continuous mechanical support to the tongue.

Also in question in the present paper are the function of tongue bracing and, relatedly, the extent to which it occurs. It has long been suggested that lingual bracing may be important in speech production for a number of reasons: to facilitate the mechanics of certain kinds of tongue movements (e.g., Stone 1990), to provide somatosensory feedback for tongue position (e.g., Stevens and Perkell 1977) and to separate the central oral tract aeroacoustically from the lateral buccal cavities (e.g., Perkell 1979,
Honda et al. 2004, 2010). Indeed, a consistent lateral seal along the tongue edge may be seen as basic to the production of any medial speech sound, as this seal, created and maintained through lateral tongue contact, creates the closed aeroacoustic tube that directs the airstream through any medial speech constriction. Notwithstanding these several reasons for supposing a priori that lateral contact should be consistently maintained throughout speech (except, minimally, during some lateral sounds where the buccal cavities are actively recruited), previous studies have generally sought bracing in only a narrow range of phonetic contexts, resulting in a widespread treatment of bracing as an occasional state, peculiar to specific sounds or sound combinations. Many examples of such studies may be found. For example, Stone (1990) refers throughout to “braced” vs. “unbraced” postures (with consonants generally falling into the former category, and vowels into the latter); Stone & Lundberg (1996) follow with three-dimensional reconstructions of articulations within the oral cavity that indicate contact against the velo-palatal periphery and alveolar ridge during selected sounds. Cheng et al. (2007) also assume that whereas vowels are unbraced, consonant articulations involve bracing, giving evidence from electropalatography of /s/ and /l/ to support the latter point. McLeod (2006) and McLeod et al. (2006) demonstrate bracing against the velo-palatal periphery for Australian English /n/ and /s, z/, respectively. Contrarily, one of several concluding hypotheses proposed by Stone (1990: 2216) speculates that it is possible that “tongue bracing is not dichotomous, but rather continuous. The familiar phonetic classes of vowels, sonorants and obstruents may roughly reflect different degrees of tongue bracing.”

The above heterogeneous view of the contexts of tongue bracing may be attributed in
part to the dominant representation of the vocal tract in midsagittal cross-section in the phonetics literature and in speech models developed for articulation-based acoustic synthesis (e.g., Rubin et al. 1981, Maeda 1990, etc.). Viewed only in the midsagittal plane, the tongue may appear to function as if unconstrained by interactions with surrounding lateral structures, more like a free-floating tentacle. Conceptualizing a more fully dimensional model of the tongue and its relationship with its surroundings in speech requires answers to questions both of basic description (e.g.: How often is the tongue in contact with surrounding structures?) and of deeper function (e.g.: What is the purpose of such contact?). Preliminary to discussions of function, however, it is necessary to determine whether contact between the tongue and other surfaces is passive or active; only if such contact is active (rather than simply being the passive result of mechanical properties of the tongue) need one inquire further as to its function.

The present paper tests the strong hypothesis that tongue bracing is both pervasive and effortful in speech, and considers whether the bulk of evidence corroborates the widely held view that tongue bracing in fact affords the tongue mechanical support. This hypothesis generates the predictions that 1) bracing should be maintained continuously throughout running speech and 2) bracing is not a passive state of the articulators, but can be shown to be achieved and maintained via active muscular control. Further to the first prediction, given the universal prevalence of speech sounds produced using a medial airstream, we predict that 1a) lateral bracing should take primacy over other types of bracing, and that 1b) in cases where lateral bracing is lost, the use of other types of bracing, such as against the alveolar ridge and lateral oropharyngeal structures, should increase.
In the remainder of this paper, evidence is drawn from articulatory data and biomechanical modeling to test these predictions. Section 2 describes a speech production study using EPG data from the Kay Palatometer Database to test the hypothesis that bracing (indicated here either by lateral contact with the sides of the dorsal palate/dentition or coronal contact with the alveolar ridge) is pervasive; we use the production data to test predictions (1a) and (1b) above. Section 3 details simulations designed to evaluate further prediction (1b), that there may be an increase in dorsal bracing that is not visible in the EPG data. Section 4 describes a simulation study designed to test prediction (2), that bracing can be shown to require active muscular control.

2. Pervasiveness of lateral tongue contact: Production study

In order to investigate the occurrence of tongue contact against the palate during running English speech, data were analyzed from the Kay Palatometer Database Model 4333. This database contains electropalatography (EPG) recordings of three spoken passages, each read by two speakers. McLeod et al. (2006) point out a perennial difficulty in using EPG for studies of lingual bracing, noting that, while incomplete lateral contact of the tongue may often been seen in EPG analysis, additional unobserved lateral contact may nevertheless be occurring via contact with the teeth. The Kay database is thus particularly valuable for studies of lingual contact as the Kay palates (no longer available) were constructed with electrodes along the teeth while the
electrodes of other EPG palates often stop well short of the teeth, as can be seen in Figure 1.

Figure 1: A Kay EPG palate (left) showing electrodes arranged all along the inner edge of the upper dentition; a more typical palate (right) built by InciDental for use with the WIN-EPG (www.articulateinstruments.co.uk) system, showing electrodes stopping well inside the molars.

Examining EPG electrode activation patterns enabled tracking of different types of palatal contact present in the database, as well as identification of any spans during which lateral contact was not observed. Of the three types of bracing under discussion in this paper, the EPG data bear on only two — lateral bracing and medial (coronal) bracing — since dorsal contact against lateral structures posterior to the pseudopalate cannot be observed using EPG.

2.1 Materials and methods

The Kay Palatometer Database includes EPG data from two speakers, one male and
one female. Dialect information for the two speakers is not available in the reference
text provided with the database, but on the basis of their productions, both speakers
apparently speak a variety of North American English.

The present study analyzed continuous EPG and audio recordings of both speakers
reading aloud the three passages included in the database: the “Grandfather passage”
(from Dworkin and Meleca 1997: 56, based on Van Riper 1963; see Reilly and Fisher
2012), the “Rainbow passage” (Fairbanks 1960), and the “North Wind and the Sun
passage” (e.g., Jones 1949), resulting in over 3 minutes of recorded continuous speech.
EPG data in the database were collected using custom-molded pseudopalates for each
speaker, each containing 96 electrodes, and were recorded using a Kay CSL 4300
machine. The resulting palatograms have a time resolution of 10 milliseconds (100 Hz).

The Kay Palatometer Database’s proprietary data-viewing software was used to display
and code EPG frames. Of the twelve horizontal rows of electrodes, the three outermost
columns of the four rows at the far back of the tongue (corresponding to the “back”
region of the palate described by Zsiga 1995) were defined as the region of interest for
lateral contact (see Figure 2). Thus, tongue contact with any of 11 electrodes on either
de side of the dorsal palate was taken as indicating contact on that side; that is, a particular
sample was considered to indicate contact on one side of the palate if at least one of
the 11 electrodes on that side in that sample indicated tongue-palate contact. Similarly,
the occurrence of coronal (anterior) contact was recorded throughout each instance of
the consonant /l/; a palatogram was considered to exhibit coronal contact if at least one
of the front-most 23 electrodes (again following regions defined by Zsiga 1995; see
Figure 2) indicated contact in that region.

Figure 2: Regions of interest defined for the Kay EPG palate based on Zsiga (1995). Shaded areas indicate electrodes used to determine lateral contact.

Using these metrics, each frame indicated either full lateral contact on both sides of the palate (bilateral contact), lateral contact on only one or the other side of the palate (unilateral contact), central contact of the anterior tongue (coronal contact), or complete loss of contact; exemplars of each type of contact are shown in Figure 3. In addition to providing examples of each of the four contact types, the annotated waveform in Figure 3 shows the utterance from which they were taken (one full phrase of the Grandfather Passage) and identifies the spans of speech during which contact was observed. We define complete (bilateral) loss of contact as a total lack of contact in the lateral target regions on both sides simultaneously (as in Figure 3d), and contrast this state with a lack of contact during which there is no contact on only one or the other side (unilateral contact, as in Figure 3c). Consequently, a frame in which there is contact against just one side of the palate is considered to be an instance of contact. This characterization
stems from the assumption that, whereas one-sided loss of contact may compromise some of the acoustic and aerodynamic correlates of bilateral contact (i.e., a complete aeroacoustic tube requires bilateral contact), any mechanical and somatosensory functions of contact could still be served by unilateral contact.

Figure 3: Example waveform, transcriptions and EPG patterns for one phrase. Gray bars above the waveform indicate time spans with coronal tongue-palate contact (“Coronal”), as well as time spans with lateral tongue contact on the left (“Left”) or right (“Right”) side of the palate. EPG frames (a, b, c, d above) were extracted at corresponding time points in the waveform indicating a) bilateral, b) coronal, c) unilateral and d) no contact, respectively.
2.2 Results

Bilateral contact was found to be sustained almost continuously throughout all recordings in the Kay database. The following subsections describe the details of our findings, first for lateral contact and then for coronal contact.

2.2.1 Lateral contact

Of the 191.7 seconds comprising all six passages in the database, complete loss of lateral contact occurred for only 4.819 seconds, meaning that lateral contact was maintained during 97.5% of the total duration of the recordings across both speakers. The female speaker exhibits a higher rate of lateral contact loss compared to the male speaker: 3.348 seconds (3.3%) versus 1.471 seconds (1.7%).

Complete loss of lateral contact was observed to occur in only two contexts: during some instances of low vowels and during some instances of the lateral consonant /l/. Among low vowels, only /ə/ exhibited widespread contact loss. The female speaker lost lateral contact during 47.6% of tokens of /ə/ (10 out of 21), while the male speaker did so during 33.3% of those (7 out of 21). For the male speaker, these were the only cases of complete contact loss during vowel articulations; all other vowels maintained contact throughout. For the female speaker, contact loss also occurred once each during articulations of three other low vowels: /æu/ (1 out of 10), /aɪ/ (1 out of 22), and /ʌ/ (1 out of 36).
Rates and allophonic conditions of contact loss during tokens of /l/ also differed somewhat between the two speakers. The female speaker's overall rate of contact loss during /l/ was 30.8% (16 out of 52), while the male speaker's was 19.6% (10 out of 51). Acknowledging known resyllabification effects (e.g., Sproat & Fujimura 1993), all prevocalic tokens of /l/ were considered to be onset allophones, while pre-consonantal and utterance-final tokens of /l/ were designated coda allophones, irrespective of word boundaries. A sharp distinction was observed between rates of complete (bilateral) contact loss for /l/ in onset position versus coda position. The female speaker exhibited complete contact loss during 17.9% of onset /l/ tokens across all passages (7 out of 39), but during 69.2% of coda /l/ tokens (9 out of 13). Similarly, the male speaker exhibited complete contact loss during 10.5% of onset /l/ tokens (4 out of 38) but during 46.2% (6 out of 13) of coda /l/ tokens. Thus, for both speakers, contact was several times more likely to be lost in coda allophones than in onset allophones of /l/, as shown in Figure 4.

Figure 4: Onset/coda asymmetries in complete lateral contact loss showing that contact is maintained in onset allophones of /l/ for both speakers.
An additional point of interest that emerges from the lateral contact data is the finding that both speakers showed an asymmetrical tendency to release contact on one side of the mouth versus the other side. This asymmetry is reflected both in the side on which contact is lost in cases of unilateral contact loss as well as in the sequential order of contact loss in cases of bilateral contact loss. Thus, the female speaker strongly favors loss of contact on the right side: among her cases of unilateral contact loss, 23 instances of loss are on the right side versus 2 on the left, while during bilateral releases, she first loses contact on the right side 21 times versus 6 times on the left side; this preponderance of right-side contact loss for the female speaker indicates a significant overrepresentation among her total instances of contact loss ($p < 0.00001$ according to a chi-square test). Conversely, the male speaker shows a preference for contact loss on the left side: unilateral contact loss for him happens on the left side 21 times as compared to only 11 times on the right side, while during bilateral releases, he first loses contact on the left side 13 times in the database versus 7 times on the right; this pattern is also significantly different from an expectation of symmetry ($p = 0.027$).

Figure 5 shows this asymmetry for both speakers.

Figure 5: Lateral asymmetry in contact loss.
2.2.2 Coronal contact

Overall, coronal contact occurs with some regularity in the database; however, of particular interest for the present paper are those cases of coronal contact that occur during periods where lateral contact has been lost. Except in cases of overlap with a coronal consonant, vowels do not generally exhibit tongue-palate contact in the coronal region, to which expectation the present database offers no exceptions. During some articulations of /l/ in this dataset, however, coronal contact was present. Our interest here is to determine whether or not coronal contact can be said to take on the role of bracing, particularly during articulations where lateral contact is lost. To this end, we measured coronal contact during articulations of /l/ specifically where lateral contact loss occurs for both speakers. Although the number of data points is small, the results suggest that coronal contact is common (occurring for 60% or 9 of 15 total tokens) when lateral contact is lost, and exceptionless (11 out of 11 total tokens) for onset allophones of /l/: for the female speaker, coronal contact occurred during 78% (7 out of 9) of coda-position /l/ tokens, and during 100% (7 out of 7) of onset-position tokens; for the male speaker, where fewer tokens were available, 33% (2 out of 6) of coda-position tokens exhibited coronal contact, while 100% (4 out of 4) of onset-position tokens exhibited coronal contact.

2.3 Speech Production Discussion

We find that lateral contact remains constant throughout nearly the entirety of all three spoken passages for both speakers, and that it is lost only under tightly constrained
circumstances: namely, a minority of low vowel and /l/ tokens. Particularly notable is the observation that, without exception, coronal contact is maintained throughout every instance where lateral contact is lost during onset allophones of /l/. Thus, we are left with a lack of EPG-measurable tongue contact in only two very limited contexts: a subset of low vowels and postvocalic allophones of /l/. Considering the many thousands of speech sounds and phonetic contexts in this corpus, the tiny number of instances of loss of contact observed in the present study is itself a remarkable finding.

To provide another view of what is happening in onset allophones of /l/, Figure 6 below shows two M-mode ultrasound contours of the first author producing the sentence “Mary had a little lamb, its fleece was white as snow” in a natural, normal-speed voice. Ultrasound imaging may be used as a tool for visualizing lateral contact (e.g., Adler-Bock et al. 2007), and M-mode ultrasound generates an image showing movement through a single intersect line plotted over time. The top track in Figure 6 shows the vertical movement trajectory occurring at the middle of the tongue dorsum — note how the midline of the tongue rises and falls dramatically with different articulations; the bottom track shows the vertical movement trajectory of the right side of the tongue dorsum, where it is in contact with the teeth. As the horizontal red lines highlight, even when the midline of the tongue is moving rapidly up and down to produce different speech sounds, the sides of the tongue remain firmly fixed, sealed against the dorsal teeth and palate. This lack of motion corresponds to the contact against the sides of the palate seen in the above EPG experiment. Only during articulations of /l/ do we see the lateral tongue pulling down, then quickly returning to its in-contact position. It is in such cases that the EPG data show the tongue tip initiating coronal contact prior to lateral
release, and releasing coronal contact only after lateral contact has once again been established.

Figure 6: M-mode ultrasound showing a constant position for one side of the tongue (“Lateral”) except for sharp dips during tokens of /l/, even while the tongue midline is moving.

Both speakers also show an asymmetrical tendency to release contact on one side of the mouth. This asymmetry is reflected both in the sequential order of contact loss in cases where it is lost on both sides, as well as in the side on which contact is lost in cases of one-sided contact loss, indicating that unilateral contact loss is best viewed as a point along a continuum between bilateral contact and bilateral loss of contact, rather than being an independent “semi-braced” state of its own.

The results presented here are consistent with a view of the tongue as remaining in almost constant contact with fixed vocal tract surfaces, either laterally or coronally, with the notable exception of a subset of instances of low vowels and postvocalic /l/. This finding prompts us to inquire further into what is special about this exceptional set. Of particular note, it has been observed that the posterior tongue postures are essentially identical for /a/ and /l/ in US English (Gick et al. 2002), with the dorsal tongue being strongly retracted into the oropharynx. In this region of the vocal tract, the posterior
tongue forms the anterior pharyngeal wall, and as such is contiguous with the lateral pharyngeal walls along its entire length, so that contact between the posterior tongue and the lateral pharyngeal walls goes without saying. Further, in the upper portion of the oropharynx, MRI evidence (Honda et al. 2010) indicates that lateral extrinsic tongue muscles such as styloglossus and possibly hyoglossus are constrained by passing through parapharyngeal tissue, further tethering the lateral tongue dorsum to the lateral pharyngeal walls. Presumably this contact only increases as the tongue is retracted into the oropharynx during low vowels. These observations give rise to a hypothesis that may explain the apparent exception to what is otherwise continuous palatal contact: That when the tongue retracts (sometimes to the point of losing lateral contact with the hard palate), the tongue dorsum increases contact with lateral oropharyngeal structures. To illustrate this hypothesis, the radiographic overlay in Figure 7 shows that retracting the tongue into the upper oropharynx appears to exert a direct rearward force on lateral pharyngeal structures (in this case, the palatopharyngeal arch). This hypothesis is further tested in the following section.
Figure 7: X-ray image showing the tongue pressing posteriorly against the palatopharyngeal arch. Black outline shows the tongue, velum, and pharyngeal structures during a particularly dorsal variant of French uvular r; white outline shows the same structures during a preceding (non-low) vowel. The X-ray image in this figure is from the Laval X-ray database (Munhall et al. 1995).

3. Pervasiveness of lateral tongue contact: Simulation study

This section describes a biomechanical simulation designed to test whether or not the dorsal tongue increases bracing contact with lateral pharyngeal structures during retracted postures for sounds such as low vowels and postvocalic /l/. Biomechanical
Simulation can be a useful tool for the analysis of vocal tract articulation in speech as it relates muscle forces, tissue mechanics, 3D shape change, and 3D movements. These biomechanical quantities are critically important to understanding the potential role of lateral tongue bracing during speech production. Simulation is useful both for interpreting experimental data and for extrapolating or interpolating where certain data are missing or very difficult to obtain. For example, simulations can help predict whether or not contact is likely to occur between the tongue and other surfaces where it is not practical to place EPG sensors. For reasons such as these, biomechanical simulations can help to clarify the overall action of the tongue observed in EPG measurements.

3.1 Methods

Our 3D finite-element tongue model is based originally on Buchaillard et al. 2009, and has been described in detail elsewhere (Stavness et al. 2012a). Importantly for the present study, it includes contact detection and resolution with surrounding structures, including the mandible, teeth, hard palate (Stavness et al. 2011) as well as the soft palate and pharyngeal structures (Gick et al. 2014). The morphology of the tongue model as well as the surrounding anatomical structures was fit to a specific speaker using 3D Computed Tomography (CT) data (Stavness et al. 2011). The tongue model includes 740 hexahedral elements (2844 degrees of freedom), and 1574 muscle fiber elements distributed throughout the mesh representing the 10 main tongue muscle groups. Material properties for the tongue are derived from experiments on human tissue (Gerard et al. 2007) and use a hyper-elastic fifth-order incompressible Mooney–Rivlin material with $c_{10} = 1037$, $c_{20} = 486$, and $c_{01} = c_{11} = c_{02} = 0$ Pa and a density of
1040 kg/m³. Incompressibility in the FEM model is implemented using a constraint-based mixed displacement-pressure formulation.

The simulations developed in this study are muscle-driven forward dynamics simulations. Muscle activations are input to the simulation, which solves continuum mechanics equations, and outputs tongue motion, 3D tongue shape, and 3D contact areas and pressures. The muscle-driven and contact-resolving simulations were performed with the open-source ArtiSynth biomechanical modeling platform (Lloyd et al. 2012). ArtiSynth, and the 3D tongue model used in this study, are freely available to download from www.artisynth.org.

Our hypothesis is that the dorsal tongue increases contact with lateral pharyngeal structures (e.g., the palatopharyngeal arch) during retracted postures for low vowels. In order to test this, low vowel tongue postures were simulated in a model that includes postlingual oropharyngeal structures, including palatoglossal and palatopharyngeal arches as well as a rear pharyngeal wall. To simulate the /a/ posture, we used previously reported muscle activations, which were determined by a procedure to qualitatively match the 3D shape of the tongue model to /a/ postures from ultrasound and magnetic resonance images (Stavness et al. 2012b). These muscle activations included genioglossus (GG), hyoglossus (HG) and verticalis (VRT) muscles (Table 1). The simulation started with the tongue at rest, ramped up muscle activations smoothly and ran until the model settled to a static-equilibrium posture. Contact regions between the tongue and surrounding structures in the final posture were depicted as cyan-colored regions on the tongue surface and their total surface area were measured.
Table 1: Muscle activations for simulated low vowel posture for /a/. Muscles include the anterior/middle portion of the genioglossus (GGA/M), hyoglossus (HG), verticalis (VRTX), and jaw opening (JAWO) muscles.

<table>
<thead>
<tr>
<th></th>
<th>GGA</th>
<th>GGM</th>
<th>HG</th>
<th>VERT</th>
<th>JAWO</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a/ Posture</td>
<td>15%</td>
<td>10%</td>
<td>15%</td>
<td>5%</td>
<td>4%</td>
</tr>
</tbody>
</table>

3.2 Results

The simulated low vowel posture resulted in no lateral contact between the tongue and upper teeth or hard palate. However, the simulation of a low vowel posture did involve a many-fold increase in contact area between the dorsolateral tongue and pharyngeal structures from 7 mm² at rest to 337 mm² at the end of the simulation. The progression of contact area for increasing muscle activation is plotted in Figure 8. Contact regions included the palatopharyngeal arch and the lateral pharyngeal wall, as illustrated in Figure 9. These results are consistent with the X-ray data in Figure 7 above showing the tongue dorsum pushing rearward into these pharyngeal structures during tongue retraction.
Figure 8: Total surface area for the regions of contact between the tongue and pharyngeal structures (dashed line) as muscle activations (solid lines) increase from rest to produce a low vowel tongue posture. Muscles include the middle/anterior portion of the genioglossus muscle (GGM/A), hyoglossus (HG) and verticalis (VRT).

Figure 9: Simulations of the jaw and tongue during production of a low vowel. Posterior (top) and lateral (bottom) 3D views of the tongue model at rest (left), at rest with pharyngeal structures cutaway (middle), and during a low vowel posture (right). Contact areas between the tongue and pharyngeal structures are shown in cyan.
4. Effortfulness of lateral bracing

The pervasiveness of lateral contact observed in the above sections still leaves unanswered the question of whether lateral contact is effortful, and hence whether it constitutes (in the usual sense of the term) bracing. Effortfulness can be examined from the perspective of whether or not lateral bracing requires active muscle effort to accomplish. That is, is the lateral tongue merely in passive contact because of its size and position in the mouth, requiring no additional effort on the part of the speaker? For example, would lateral /l/ require a speaker to pull the sides of the tongue actively away from the teeth? Or alternatively, is lateral tongue contact actively maintained through extra muscular effort, so that lateral /l/ is achieved, at least in part, through the relaxation of tongue muscles? Biomechanical modeling is well suited to answer questions related to muscle effort because it can relate changes in the shape and contacts of the tongue to changes in muscle forces.

4.1 Methods

We used muscle-driven tongue simulations to assess the muscular effort required to establish lateral contact. Muscle inputs to such simulations may be informed by electromyography measurement (Sartori et al. 2014), but reliable validated electrophysiological data for the tongue are exceedingly sparse and difficult to obtain. Alternative simulation schemes such as tracking-based inverse simulations are able to predict muscle activations (Stavness et al. 2012a), but are challenging to implement,
and thus far only allow flesh-point targets rather than contact area (e.g., tongue-palate) targets. Further, given the large number of tongue muscles it may be possible for different combinations of muscle activations to result in similar tongue position, shape, and/or contacts with the palate. Manual tuning of muscle activation inputs is commonly employed in simulation studies, but even with expert knowledge and careful consideration, care must be taken in order to not produce biased simulations. To avoid bias when generating simulations of lateral tongue contact, the present paper employed a large-scale set of simulations testing all possible combinations of muscle recruitment in the tongue model with three activation levels (0%, 20%, 50%). These combinations resulted in approximately 60,000 simulations from which it is possible to derive which muscle groups are consistently associated with particular patterns of tongue-palate contact.

To interpret the simulation outputs pertaining to questions of lateral tongue bracing, we added a virtual EPG analysis to the model, whereby virtual contact sensors were placed on the hard palate and upper teeth of the model in locations consistent with the Kay EPG (Figure 10). The virtual EPG analysis allowed us to categorize each of the 60,000 tongue simulations into the same contact regions of interest as were used for the experimental EPG data in Section 2.
In order to determine if active muscular effort is required to establish lateral contact during medial tongue contact (as in a coronal stop consonant), we used the virtual EPG data to find the subset of all simulations, out of all 60,000 muscle combinations, that resulted in medial contact without lateral contact (unbraced) and medial contact simultaneous with bilateral contact (braced). We expect that cases of contact that are more rare must require more precisely controlled muscle activations. In addition to
frequency of occurrence, we also identified the simulations with minimum muscle effort within the braced and unbraced sets.

Muscular effort for each simulation was measured as the sum of squared muscle activations:

\[ \text{Effort} = \sum_{i=1}^{n} a_i^2 \]

where \( n \) is the number of muscle groups and \( a \) the muscle activation level between 0.0 – 1.0. This is a common criterion used in optimization studies as a surrogate measure for metabolic energy expenditure (Erdemir et al. 2007, Stavness et al. 2012a).

We visualized the model for the minimum muscle effort simulations of the braced and unbraced sets to qualitatively evaluate the 3D tongue shape and medial contact region.

4.2 Results

Our simulation results indicated that out of 60,000 different muscle combinations, only 1000 resulted in bilateral contact. Bilateral contact generally required strong activation of superior longitudinal (SL), posterior/middle genioglossus (GGP/M), mylohyoid (ML), and verticalis (VRT) and low activation of inferior longitudinal (IL), anterior genioglossus (GGA) and hyoglossus (HG) muscles.

We found 220 (out of 60,000) muscle combinations that produced medial contact, but only 81 muscle combinations that could produce both medial contact and bilateral
contact simultaneously. This suggests that maintaining lateral contact while creating a medial tongue constriction requires a more precise specification of muscle activations, whereas producing a medial constriction alone permits a wider range of muscle activation combinations.

Muscle activations for the simulations that achieved medial tongue-palate contact with the minimum muscle effort are plotted in Figure 11. The two simulations achieved approximately the same medial contact, but with different overall 3D tongue shapes (Figure 12). The minimum muscle efforts required for braced and unbraced medial contact were 0.33 and 0.12 respectively, supporting the view that producing a medial tongue-palate constriction with lateral contact requires additional muscle effort as compared to the same constriction without lateral contact.

![Figure 11: Muscle activations for simulations of medial tongue tip contact with and without lateral bracing. Muscles include the posterior/middle/anterior portion of the genioglossus muscle (GGP/M/A), as well as the styloglossus (STY), mylohyoid (MH), hyoglossus (HG), verticalis (VRT), transversus (TRN), and inferior/superior longitudinal (I/SL) muscles.](image-url)
Figure 12: Simulations of medial tongue contact with and without lateral bracing. Midsagittal (top) and coronal (bottom) cutaway sections of the model showing the virtual EPG sensors (colored spheres). Sensors in contact with the tongue model are highlighted and marked with arrows.

While there may be many possible mechanisms for effecting lateral bracing, it is clear that lateral bracing requires additional muscle recruitment and consequently expends additional energy, indicating that lateral bracing is an effortful process.
5. Discussion

Our initial hypothesis maintained that bracing of the tongue is both pervasive and effortful in speech. We have provided evidence for the validity of two general predictions generated by this hypothesis, namely that 1) bracing is maintained continuously throughout running speech, and 2) bracing requires active muscular control. The first prediction breaks down into at least two sub-predictions: 1a) lateral bracing occurs throughout running speech except when released for specific sounds, and 1b) during those sounds where lateral bracing is lost, bracing in other locations is increased. Prediction 1a is supported in the present paper by EPG results, prediction 1b by a combination of production and simulation results, and prediction 2 by simulation results. The simulation study in section 4 implicated the following functional contributions of muscle actions in lateral tongue bracing: Verticalis (VRT) flattens the tongue body and consequently expands the tongue body laterally (due to the hydrostatic nature of tongue tissue) and mylohyoid (MH) raises the tongue body so that the lateral margins are at the height of the upper teeth; in addition, genioglossus posterior (GGP) and superior longitudinal (SL) provide additional activation needed to maintain anterior contact.

Prediction 1b, that bracing is increased in other locations when lateral bracing is lost, is particularly compelling in establishing bracing (in general) as a mandatory and active aspect of speech production. As the observations supporting this prediction concern primarily /l/, the results of the present study have the unexpected benefit of providing a unified account for several previously unexplained aspects of /l/ production. Thus, for example, when the sides of the tongue are pulled away from the teeth to produce lateral
airflow, one possible explanation could be that the tongue simply elongates, resulting in passive coronal tongue-palate contact for lateral sounds such as /l/; if this were the case, we should expect coronal tongue-palate contact to occur subsequent to lateral bracing loss in onset /l/ allophones. However, the present EPG results show that coronal tongue-palate contact precedes loss of lateral bracing, so that coronal bracing is always in place before lateral bracing is lost. This sequencing indicates an active effort to ensure that bracing is always maintained, corroborating the view that bracing plays an essential mechanical role in the speech production system. Similarly, this account provides an independent and principled explanation for known interarticulator timing patterns previously seen in /l/ allophones (e.g., Sproat & Fujimura 1993, Gick 2003, Gick et al. 2006, Scobbie & Pouplier 2010) wherein the tongue anterior constriction precedes tongue dorsum backing in onset allophones, as well as the observation that tongue-palate contact generally occurs during onset “bright” but not coda “dark” allophones of /l/ (Giles & Moll 1975).

This paper also contributes to our understanding of lateral asymmetry in lingual motor behavior. The finding that both speakers in the EPG study exhibited a significant tendency to release bracing on one side of the tongue rather than the other mirrors the findings of McAuliffe et al. (2001) concerning asymmetries in the production of English [s], [t], and [k]. While McAuliffe et al. (2001) found no significant lateral asymmetries in [l] production, their study did not investigate contact at the edges of the tongue. We find it unsurprising that lateral asymmetries in [l] emerge only at the tongue edges, since a defining characteristic of [l] is the loss and re-establishment of lateral contact. Our finding that the two EPG study speakers differ from each other in the
directionality of their lateral asymmetries also suggests a possible connection with handedness or other lateral asymmetries in the body (cf. Pribram 1977), a topic warranting further study.

Simulation studies face a number of common challenges, which we addressed through multiple safeguards in the present study. One such challenge is experimenter bias. For example, manual selection of model parameters could be tuned or conveniently chosen to elicit an intended result. To avoid this kind of bias, the model of the vocal tract used in the present study was built to model as closely as possible the properties of the body for other purposes (surgical modeling, swallowing biomechanics, etc.) without reference to the present theoretical questions. Also, within this larger model context, the present study used a novel simulation methodology whereby uniform sampling of a large number of muscle activation combinations was performed to explore a variety of muscle-driven tongue configurations. This sampling approach counters experimenter bias in that it permits analysis of specific configurations, e.g. lateral tongue bracing, without allowing bias in the selection of muscle activations. Another challenge of simulation studies is that some model outputs may be over-sensitive to small changes to input parameters. Our uniform sampling methodology permits an analysis of the sensitivity of the configuration of interest to the parameters sampled (muscle activations, in our case). In addition, for simulations of pharyngeal contact our results show that contact area is not overly sensitive to muscle activation magnitude. Finally, the present simulation study may potentially lack generalizability as the model is based on the morphology of a single speaker. Our team is developing tools to adapt models to other speakers (Harandi et al. 2016), which will permit us to examine generalizability
across speakers in future studies. Despite these potential limitations, this simulation study represents, to our knowledge, the largest dataset of 3D biomechanical tongue simulations ever reported. We have made the model, simulator and simulation data publicly available (http://www.artisynth.org).

The finding that contact between the tongue and peripheral structures in the oral cavity is neither incidental nor passive, but rather a pervasive and effortful feature of speech articulation, challenges prevalent conceptions of tongue motor control. Rather than the tongue functioning as a free-floating, tentacle-like structure moving as a mass into more or less high-low or front-back positions through the action of extrinsic muscles, the tongue in this view remains by and large stable, with constrictions formed largely through changes in shape effected by the actions of intrinsic muscles (see, e.g., Honda et al. 2013). Lingual constrictions in this view are more 3-dimensional and sphincteric (or, in the case of palatal constrictions, hemi-sphincteric), more straightforwardly analogous to constrictions elsewhere in the vocal tract (e.g., lips, larynx, pharynx, velopharyngeal port), and easier to relate to oral mechanisms associated with other (digestive, respiratory) functions of the vocal tract. We hope that a better understanding of tongue bracing along with the concomitant advances in computational modeling of speech biomechanics (e.g., Stavness et al. 2012b, Gick et al. 2014) will enable broader applications, including improved training for clinicians (McLeod 2011) and more grounded and embodied theories of speech production (Gick & Stavness 2013).
Acknowledgements

The authors thank S. Bird, C. Chiu, D. Erickson, S. S. Fels, A. Klenin, T. Magnuson, S. Moisik, M. Stone, D. H. Whalen, I. Wilson, and N. Yamane for their various contributions to this work. We acknowledge funding from NSERC Discovery Grants to the first and last authors.
References


