

Lateral tongue bracing as a universal postural basis for speech

Yadong Liu 

University of British Columbia
yadong@alumni.ubc.ca

Felicia Tong

University of British Columbia
feliciatong@yahoo.com.hk

Gillian de Boer

University of British Columbia
gillian.deboer@ubc.ca

Bryan Gick

University of British Columbia & Haskins Laboratories
gick@mail.ubc.ca

Lateral bracing refers to an intentional tongue posture whereby the sides of the tongue make contact with the sides of the palate and the upper molars. While previous research on this topic has focused mostly on English, the present study tests the hypothesis that lateral bracing provides a fundamental postural basis for speech and is present across languages. We predicted that, across multiple languages, the sides of the tongue should be more stable than the center and should stay in a relatively high position in the mouth throughout most of running speech. Using coronal ultrasound imaging, we measured tongue movement produced by speakers ($N = 28$) of six languages (Akan, Cantonese, English, Korean, Mandarin and Spanish). Across these languages, as predicted, the sides of the tongue throughout running speech were positioned higher in the mouth than the center, and the range of movement of the sides was significantly smaller than that of the center of the tongue. These findings support the view that the sides of the tongue maintain a braced posture across languages while speaking, potentially constituting a universal, rather than language-specific, postural basis for speech.

1 Introduction

Tongue BRACING is a term used to describe postures in which the tongue is held in contact with surrounding hard surfaces, including the palate and/or teeth (e.g. Stone 1990). LATERAL TONGUE BRACING specifies a posture where the sides the tongue are held in contact with the edges of the palate and upper molars in the posterior oral cavity (as opposed to medial bracing, for example, in which the tongue makes contact with the center of the palate, alveolar

ridge and/or upper incisors, as in some instances of /l/; see Gick, Wilson & Derrick 2013). The human vocal tract can be viewed as an aeroacoustic tube, where the anterior portion of the tube is formed by the palate above and the tongue below, with contact between the two along the sides (Gick et al. 2017). A laterally braced tongue posture contains airflow and resonance within a central cavity, so that changing the shape of that cavity can produce a variety of different speech sounds. Stevens & Perkell (1977) showed that tongue bracing also provides somatosensory feedback for tongue position, which is crucial for speaker's speech perception as such feedback can affect auditory identification (D'Ausilio et al. 2014). If this view of tongue bracing as a fundamental postural basis for speech is correct, then bracing should be a pervasive feature of every language in the world. While the existing literature has shown evidence of pervasive bracing in English (Bressman et al. 2010, Gick et al. 2017), studies to date have yet to establish whether this pervasive bracing posture is maintained across other languages.

Previous research has documented tongue–palate contact in speech, often with a focus on specific speech sounds. Many of these studies have focused on English, and most have used palatography and electropalatography (EPG). For example, McLeod (2006) and McLeod, Roberts & Sita (2006) used EPG to show lateral bracing during the productions of /s/, /z/ and /n/ for Australian English speakers. Another EPG study (H. Cheng et al. 2007) found that bracing was used by English-speaking children as well as adults. Gick et al.'s (2017) study of archival EPG data on English found that lateral tongue bracing was maintained throughout more than 97% of continuous speech. Bracing was only disrupted for some instances of laterals and low vowels such as /a/, /ʌ/, /aʊ/ and /aɪ/, and the overall rate for contact loss was 25.2% (26 out of 103) during laterals and 40.5% (17 out of 42) during /a/ combining both the female and the male speaker (Gick et al. 2017). The rate for contact loss was 10% (one out of 10) during /aʊ/, 4.5% (one out of 22) during /aɪ/, and 2.8% (one out of 36) during /ʌ/ for the female speaker (Gick et al. 2017). No contact loss was observed for /aʊ/, /aɪ/ and /ʌ/ during the male speaker's speech (Gick et al. 2017). Though lateral bracing has also been demonstrated in English in phrase-level and continuous speech (Bressmann et al. 2010, Gick et al. 2017, Lulich & Pearson 2019), and has been observed to be pervasive across nearly all sounds (Bressmann et al. 2010, Gick et al. 2017), bracing has only been demonstrated for other languages at the level of individual speech sounds.

Some previous studies have reported lateral bracing in specific speech sounds of non-English languages. For example, Kochetov & Colantoni (2011) investigated articulatory patterns of coronal consonants in Cuban and Argentine Spanish using EPG. In their work, lateral contact for various consonants corresponding to the orthographic <t ch n ñ s z ll y l r> was observed. Their work also suggested that less side contact was observed for the lateral compared to the trill, and complete loss of lateral contact at one side was observed for /a/ in *arar*. (Kochetov & Colantoni 2011: 328). Kwok & Stokes (1997) examined tongue–palate contact patterns of intervocalic Cantonese consonants including /s t t^h ts ts^h l n k k^h kw k^hw ŋ/. Complete loss of lateral contact was observed for some instances of /l/ and /a/, the /a/ being part of the carrier syllable /aCa/ (Kwok & Stokes 1997). H. Kim (2001) investigated tongue–palate contact patterns of Korean obstruents using palatography, and found lateral contact for /t s c/. D. W. Kim (2000) examined lingual gestures of allophones of the Korean liquid phoneme (which the author referred to as tap /r/, and EPG results suggest that more lateral contact was observed for the syllable-initial tap [r] than the lateralized word-final [l]). Investigation of tongue–palate contact patterns of Mandarin consonants showed that lateral contact was observed for /t t^h s ts ts^h tʂ tʂ^h n/, and loss of contact was observed for /l/ (Yu, Duan & Zhang 2013). Corneau (2000) observed lateral bracing for French /t d t+j d+j/ and high vowels /i u y/, and more lateral contact was observed when the consonants were followed by high vowels compared to by low vowels. These findings are similar to Gibbon et al.'s (2007) EPG results, where more lateral contact was observed for English alveolar

stops when followed by a high vowel compared to by a low vowel. Again, note that all of the above non-English EPG studies are limited to observing tongue bracing in specific phonemes or sets of phonemes.

One challenge to using EPG is that custom pseudopalates need to be built for each participant, making it difficult to collect data from a wide range of speakers. In addition, EPG arrays often lack sensor points which can detect tongue–upper-teeth contact (Gick et al. 2017), making it difficult to assess lingual bracing against the posterior teeth and lateral palate. In the current study, we propose to test the presence of lateral bracing using ultrasound imaging. Previous studies (of English only) have occasionally observed lateral bracing using lingual imaging approaches. Stone (1990) examined English speakers' tongue movement using ultrasound (coupled with X-ray microbeam), and found some lateral tongue bracing during vowel production. Using coronal ultrasound imaging, Bressmann et al. (2010) found that English speakers had a greater amount of movement for the center of the tongue than the sides, proposing that the observation supported lateral bracing. More recently, Lulich & Pearson (2019) examined the production of an English phrase by superimposing ultrasound coronal tongue images with the equivalent MRI coronal plane image of the articulations. After alignment they were able to demonstrate lateral tongue bracing for /t/ with the sides of the tongue making contact with the palate and upper teeth (Lulich & Pearson 2019).

In the only previous cross-language study of tongue bracing, Cheng, Schellenberg & Gick (2017) presented data from a pilot study using coronal ultrasound imaging to observe tongue movements of six native speakers representing six different languages: Cantonese, Korean, Mandarin, Portuguese, Spanish and Turkish. They asked participants to read the passage 'The North Wind and the Sun' in their native language, and found that the sides of the tongue stayed predominantly in a higher vertical position with minimal movement. Their results suggested that, during the majority of their running speech, these individual speakers were holding the sides of their tongues in a braced position. However, Cheng et al.'s (2017) findings are difficult to interpret because of methodological limitations. Hand tracing coronal ultrasound images is somewhat subjective. Also, tongue–palate contact in ultrasound images can only be inferred, not seen, so an arbitrary vertical location was used as a threshold to define tongue–palate contact in their study. Tongue–palate contact was reported when the tongue stayed above the threshold. Had a different threshold been chosen, it would have resulted in a different proportion of time the tongue was reported to be in braced position. Furthermore, while tongue–palate contact of the sides of the tongue was examined, the height of the center of the tongue was not reported, leaving it unclear whether there was a difference between the positions of the center and the sides of the tongue for these speakers.

The current study is intended to test whether bracing occurs across languages by measuring tongue positions in Akan, Cantonese, English, Korean, Mandarin, and Spanish. It aims to build on the work of Cheng et al. (2017) by investigating whether lateral bracing is maintained cross-linguistically while addressing some of the previous study's limitations. In this study, data from a larger sample of participants per language were collected, and data analysis was conducted using a more automated procedure, allowing for more reliable and replicable measurements. Moreover, a video analysis was conducted to validate the presence of tongue–palate contact. Intraoral videos showing tongue–palate contact were taken via a camera and ultrasound imaging showing coronal tongue slice was recorded simultaneously. Lastly, the study was designed to measure and compare movements of both the sides and the center of the tongue. If our hypothesis that lateral tongue bracing is a universal posture for speech is correct, we predict that, for all languages, (i) the movement of the sides of the tongue is generally restricted to a smaller region than the center of the tongue in the posterior oral cavity, (ii) the sides of the tongue are situated in a higher position in the mouth than the center of the tongue in the posterior oral cavity, and (iii) the sides of the tongue are frequently in contact with the roof of the mouth in the posterior oral cavity.

2 Method

As per Gick et al. (2017), lateral bracing occurs when the edges of the back of the tongue contact the lateral back region of the palate, as illustrated in the left panel of Figure 1. A posterior coronal view of the tongue–palate lateral contact is shown in the right panel of Figure 1. Ultrasound was used to collect coronal plane images of the tongue dorsum while participants read a passage in their native language. We then conducted a novel analysis of the vertical position of this portion of the tongue to investigate the presence of lateral bracing (see Section 2.2). Ethics approval for this study was granted by the Research Ethics Board of the University of British Columbia.

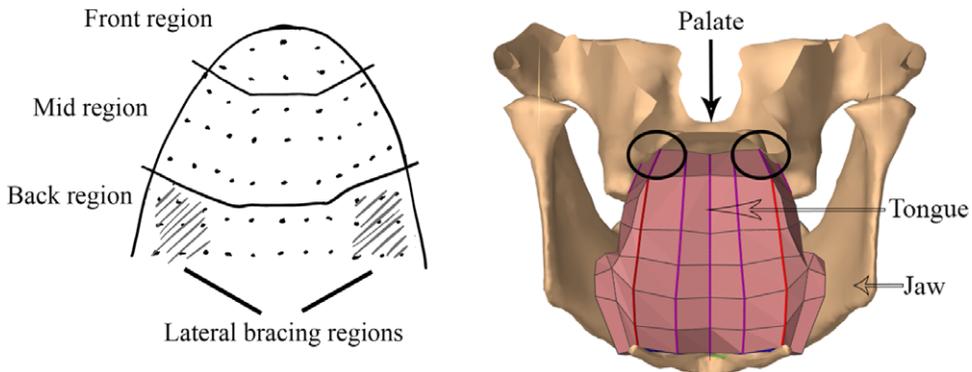


Figure 1 (Colour online) The left image shows palate regions, including lateral bracing regions, after Zsiga (1995). The right image (generated using ArtiSynth <www.artisynth.org>) shows a cutaway posterior view of the tongue held in a laterally braced position against the palate as circled.

2.1 Data collection

Data were collected from 80 speakers of six different languages: Akan, Cantonese, English, Korean, Mandarin, and Spanish. Akan belongs to the Kwa group of the Niger-Congo family (Pulleyblank 2018). It was chosen because it belongs to a separate language family from the other five languages and, in addition, Akan's phonemic inventory contains no lateral sounds (Dolphyne 1988). Along with low vowels, lateral sounds have been identified as the primary source of interruption of lateral bracing in English (Gick et al. 2017), and it is unknown whether bracing behaves differently in a language without lateral sounds. While both Cantonese and Mandarin belong to the Sino-Tibetan language family (Dryer & Harnsperger 2013), they differ markedly in lingual articulatory gestures – and therefore possibly in bracing as well – as Mandarin has alveopalatal and retroflex series (Lee & Zee 2003) that are absent in Cantonese (Zee 1991). Canadian English was chosen to provide a baseline, as lateral bracing in English has been investigated in many previous studies. Korean was chosen because it is an isolated language, thus lacks relation to the other chosen languages (Dryer & Harnsperger 2013, Song 2006). Also, Korean laterals behave differently than in other languages as the lateral /l/ is realized as [l̥] at syllable-final position, and as [r̥] at syllable initial position (Lee 1993, Kim 2000, Ha, Johnson & Kuehn 2009). Lastly, Mexican Spanish was selected as an Indo-European language that has not been previously studied for lateral bracing in running speech and has a markedly different sound inventory from that of English. For instance, denti-alveolar trill and velar fricative are found in Mexican Spanish but not in English.

Participants were considered to be native speakers if they acquired the language before five years of age and continued to use it as a primary language at work, school and/or home.

As the present study tests the hypothesis that bracing is universal, more detailed information about dialects or variants of languages was not recorded. The majority of the participants were recruited through an anonymous participant pool, such that data had to be collected irrespective of language background. There were 26 speakers of English, 14 Mandarin, nine French, eight Akan, eight Korean, eight Cantonese and seven Spanish. Data collection was stopped when there were at least seven speakers for each language. Due to poor image quality (see below), some speakers' datasets had to be discarded. We also excluded data from participants who reported being native speakers but did not demonstrate native-like proficiency as they struggled with the reading task; this was particularly the case for some of the participants who reported speaking French as a native language. Four to five speakers per language were retained and analyzed. Where there were more than five usable datasets for a language (e. g. English), the first five were selected. For the speakers of French, between image quality and concerns about the speakers' familiarity with the language, there were less than four usable datasets and so the language group was put to the side. In the end, data from 28 speakers were analyzed (mean age 23.5 years, 10 males).

Each participant was asked to read aloud a translation of the passage 'The North Wind and the Sun' (IPA 1999). All participants could read the translation of the passage fluently. Translation of the passage is readily available for Spanish (Avelino 2018), Cantonese (Zee 1991), Mandarin (Lee & Zee 2003), and Korean (Lee 1993). The passage was translated to Akan by Wendy Amoako, who is a trained linguist and speaks both English and Akan fluently. Use of this standard passage ensured that speakers would produce most of the phonemes of the language in running speech, though it did not allow for controlled balancing of phoneme distribution across languages. Sound inventory of Akan was obtained from Amoako (2020), and the inventory of other languages was obtained from the Illustration of the IPA including Cantonese (Zee 1991), Korean (Lee 1993), English (Hillenbrand 2003), Mandarin (Lee & Zee 2003) and Spanish (Avelino 2018). Missing phonemes are /g/ in Akan, /p^h k^{wh} u ɔi ui iu/ in Cantonese, /dʒ ɔɪ/ in English, /k^h je je ja jo ju we wɔ uui/ in Korean, /ts^h tɕ^h w y iu uai/ in Mandarin, and /ɲ j əj əj eɪ eɪ oɪ uɪ ju wi wo/ in Spanish. Moreover, as producing laterals and low vowels have been reported to disrupt bracing (Gick et al. 2017), the proportion of laterals and low vowels in the passage 'The North Wind and the Sun' is shown in Table 1. The latter includes monophthongs in the lower third of the vowel diagram and diphthongs that contain at least one low vocoid.

Table 1 Count of lateral, low vowels and the proportion of such phonemes in the passage 'The North Wind and the Sun' in different languages.

Language	Low vowel: Lateral count							Total count of laterals and low vowels/segments	Ratio
Akan	a: 61	æ: 8						69/433	0.16
Cantonese	a: 7	ɛ: 20	au: 2	ɛu: 11	ai: 5	vi: 6	l: 17	68/314	0.22
English	e: 5	æ: 6	ʌ: 7	ɑ: 6	aɪ: 2	au: 1	l: 18	45/373	0.12
Korean	a: 28	ʌ: 17	ja: 0	jʌ: 8	wa: 2	wʌ: 0	l: 12	67/367	0.18
Mandarin	a: 16	ai: 11	ia: 7	au: 8	ua: 4	l: 20		66/345	0.19
Spanish	a: 47	aj: 0	aj: 0	ja: 5	wa: 5	l: 29		86/408	0.21

The first transcribed sentence from the passage in each language is shown below (laterals and low vowels are underlined):

Akan

[æ̣tífí !ṃfiámá ní òuǰiá dzíí æ̣tɛ́ɲé wò wò mú diè nì hò yé dǐ, bíí á ðk^wáǐtúní bí á ódí ntòmá
diùdiù ákátá nì hǒ bé téeé mú]

Cantonese

[jau̯1 jət̯ ts^hi1 lɛ̯ pək̯1 fɔŋ1 t^hoŋ1 t^hai̯1 jœŋ1 hɛ̯i1 tou̯1 au̯1 kɛ̯n1 pin1 kɔ̯1 lɛk̯1 ti1 wɔ̯1 k^hœy1 tei1
am̯1 am̯1 t^hɛ̯i1 tou̯1 jɛ̯u̯1 kɔ̯1 jɛ̯n1 haŋ1 k^wɔ̯1 li1 kɔ̯1 jɛ̯n1 tsæk̯1 tsy1 kin1 tai̯1 lɛ̯u̯1]

English

[ðə 'no:θ ,wɪnd ən ə 'sʌn wə dɪs 'pjʊrŋ 'wɪtʃ wəz ðə 'stʊŋgə, wɛn ə 'tɹævlə ,kɛm ə 'lɑŋ
'ræpt ɪn ə 'woɪm 'kɪlɔk]

Korean

[ba̯'ramgwa 'hennimi zʌ'ro hi'mi 'dʌ 'ze:dəgo da̯'t^hugo i'suɪ te, 'hʌn na̯'gʊnɛga ta̯'tut^hʌn
'we:t^huruɪ 'ibko 'gə:rʌ wʌ'sumnida̯]

Mandarin

[iəu̯1 i1 xuei̯1 pei̯1 fəŋ1 kən1 thai̯1 iəŋ(J) tʃəŋ1 tsai̯1 naə̯1 tʃəŋ1 luən̯1 fei̯1 də(1) pən̯1 liŋ1 ta̯1
fuo̯1 dʒə(J) fuo̯1 dʒə(J) lai̯1 lə(1) i1 kə(J) kuə̯1 lu̯1 tə(J) fən̯1 fən̯(J) tʃ^huan̯1 lə(J) i1 teien̯1 xou̯1
p^hau̯1 tsɿ̯1]

Spanish

[eɫ 'bjento deɫ 'norte jeɫ 'soɫ es 'tabʌn disku'jendo sobre 'kwa de los 'dos era eɫ 'mas 'fwerte,
kwando pa'so umbja 'xerom 'bweɫto enuna 'grwesa 'kap^ha]

The passage was read three times in the speaker's native language while a coronal section of the back of their tongue was imaged using ultrasound. Separate ultrasound imaging recording was collected for each entire passage at the resolution of 640 x 480. Pauses at the beginning (before starting passage reading) and end (after finishing passage reading) of each recording were excluded. The participants were asked to sit in a chair and lean their head against two pad supports firmly attached to the chair to keep the head relatively stabilized. A magic arm attached to the same chair held the ultrasound transducer in a fixed position to the skull for a coronal section scan of speakers' tongue surface. To capture the bracing movement, an Aloka ProSound 5000 ultrasound probe was positioned under the posterior part of the chin and aimed towards their upper molars. To determine the angle of the probe, participants were asked to repeat 'Mary had a little lamb, little lamb, little lamb'. The position and angle of the ultrasound probe was adjusted until the sides of the tongue were seen to release during production of the laterals in 'little lamb'. This sentence was chosen because it contains several laterals where the lateral tongue is most likely to be released from a braced position (Gick et al. 2017). Hence, this sentence helps to locate the lateral bracing region

shown in Figure 1 above as repetitive lateral releases were only observed in this region when producing ‘little lamb’ (Gick et al. 2017). The entire experiment for each participant was about 25 minutes.

Ultrasound images produced by the system were transferred synchronously to a computer and recorded as movie clips using iMovie. From the three repetitions collected, the repetition with the clearest image (where the tongue surface contour was most prominent throughout the entire passage), was selected for further analysis. Substantial individual variation in image quality was also observed. A speaker’s data were excluded if a clear tongue surface was not presented consistently in any of the repetitions. One factor contributing to poor image quality when imaging the lateral tongue is that, depending on tongue morphology, the ultrasound signal may intersect with the lateral sublingual air cavity beneath the edge of the tongue (La’Porte, Juttla & Lingam 2011), similar to the effect of the anterior sublingual cavity obscuring the tongue tip in midsagittal ultrasound imaging (Lundberg & Stone 1999). Moreover, the sides of the tongue are more likely to be obscured when the tongue is in a high position (Stone 2005), which we predict for the braced posture.

2.2 Data analysis

Vertical tongue positions and their distributions were used as indicators of the behavior of the tongue dorsum during continuous speech. A smaller range of movement of the edges of the tongue relative to the center indicated that the sides of the tongue were positionally more stable than the center. A distribution with a higher or lower vertical position would mean that that portion of the tongue tended to remain in a higher or lower region, respectively. A skewed vertical distribution would indicate a mechanical constraint such as meeting a hard structure (e.g. the roof or floor of the mouth). Thus, bracing of the tongue against the roof of the mouth would be consistent with the sides of the tongue having (i) a reduced range of movement, (ii) a relatively high position in the mouth, and (iii) a skewed vertical distribution. This method was validated against data synchronously collected by a camera directly showing contact between the sides of the tongue and the roof of the mouth. Further details on validation will be discussed in Section 2.3.

To test the three above measures of bracing, we identified the range of motion and relative height of the sides and the center of the tongue during production. The ultrasound images were then analyzed by tracking vertical movement using videokymography (henceforth VKG), a video analysis method originally developed for plotting vocal fold movement over time (Švec & Schutte 1996). An example is illustrated in Figure 2. The image on the left in Figure 2 shows one coronal image of the tongue with a vertical slice (in blue) through the right side of the tongue. This image was converted to black and white for illustration purposes. The center image (enlarged) represents slices of each frame being extracted, with the bottom pixel (red) representing the surface of the tongue moving over time, from one frame

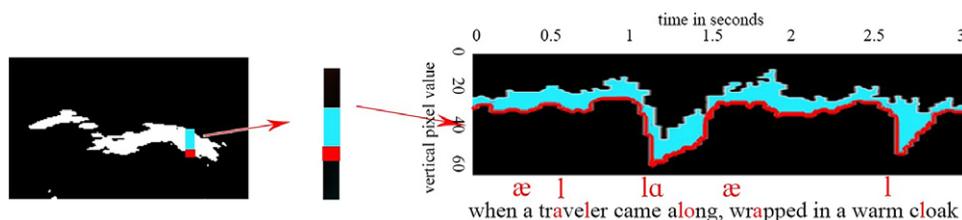


Figure 2 (Colour online) An ultrasound image on the left, with a vertical section from the right side of the tongue being measured repeatedly over time (shown as enlarged in the center image in blue with the bottom pixel in red), and the resulting ultrasound VKG on the right. Panel on the right represents the VKG showing the vertical movement of the right side of the tongue during the production of an English sentence fragment *when a traveler came along, wrapped in a warm cloak*. Utterance of laterals and low vowels have been annotated.

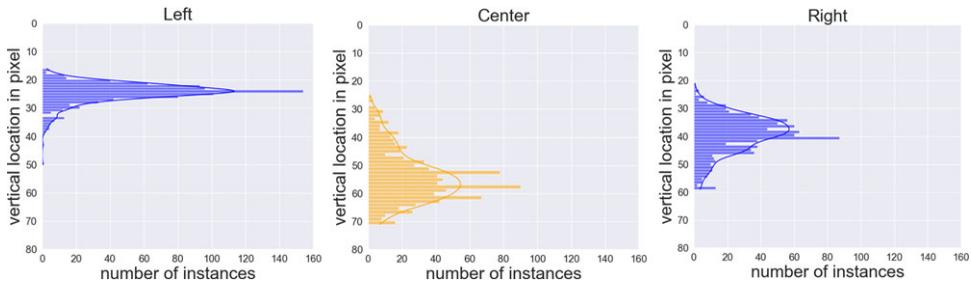


Figure 3 (Colour online) A sample distribution of one speaker's vertical position of the tongue in three regions of interest: left (blue), center (orange), and right (blue). The y-axis shows the span of vertical location (pixel values 0 from top) from the ultrasound VKG; the x-axis shows the number of instances at each vertical pixel value. These distributions show that the sides of the tongue stay high and within a small area, while the center of the tongue moves in a larger area at a lower location.

to the next. The image on the right in Figure 2 is an ultrasound VKG, produced using ImageJ 1.51 (Schneider, Rasband & Eliceiri 2012). Each grayscale ultrasound VKG was converted to black and white by applying image thresholding in ImageJ 1.51 (Schneider et al. 2012) so that the tongue position over time is clearly visible and surrounded by no speckle noise (Stone 2005). The same threshold was applied to all three regions for the same speaker. Further, speckle noise appearing elsewhere in the ultrasound VKG (i.e. not close to the tongue tracing) was removed. The VKG is composed of the same vertical section sampled over time, and the red contour depicted in the ultrasound VKG (Figure 2) shows the tongue surface's vertical movement over time. The ImageJ software assigned a pixel value of zero to the top of the image, thus the red instances located higher in the image are assigned a smaller pixel value.

Figure 3 shows sample distributions of the tongue surface position in the three regions of interest: the left, center, and right. These regions were selected because, at least in English, lateral bracing is maintained by the sides of the tongue making contact with the molars and lateral palate, while the center of the tongue moves to produce various sounds (Gick et al. 2017). Since the visibility of the tongue surface in ultrasound can be variable, there were instances when a pixel value was missing for the center, right and/or left of the tongue in a particular frame. As such, there will be subtle variations among the number of reported pixel heights for each region of the tongue.

The intersextile range (ISR; i.e. the region between the highest and lowest sextiles rounded down, or middle 68 percent of the distribution) of the tongue vertical position was used as a measure of where the tongue stays most of the time during production (qualitatively similar patterns of results were observed for other center ranges of the data, e.g. interquartile range, interoctile range). A wider range of vertical positions corresponds to a larger ISR value. The ISRs of the three regions of interest were compared across speakers and languages using a Wilcoxon signed rank test, as normal distribution was not assumed for the ISR of regions across speakers. If the ISRs of the sides of the tongue are significantly smaller than the ISR of the center of the tongue across speakers, this would indicate that the sides undergo less vertical movement and are more stable than the center of the tongue.

Where different regions of the tongue stay most of the time was identified by calculating the median and the skewness of the distribution. For example, a positive high-skewed distribution indicates that the frequently occupied area is centered at a smaller pixel value (along the y-axis) than it would be if the distribution were not skewed, which in turn shows that most of the time that region of the tongue is located at a relatively high vertical position. Whether the sides of the tongue stay higher than the center is determined by the median of the distribution. For instance, a region of the tongue that has smaller median values than others would indicate that it stays in a higher vertical position than other regions.

The direction of skew was tested using Shapiro–Wilk tests for each speaker. If the distributions of the vertical regions of the sides of the tongue are significantly high-skewed, while the distributions of the center of the tongue are less so, this would indicate that the sides of the tongue are generally remaining in a relatively high vertical position whereas the center of the tongue stays at a relatively low vertical position most of the time. All statistical tests were carried out in R version 3.6.1 (R Core Team 2019).

2.3 Validation

Video analysis was used to validate the interpretation of the above ultrasound imaging analysis regarding the sides of the tongue staying higher and making contact with the roof of the mouth. Two English-speaking participants were asked to read aloud a one-minute-long passage (see Appendix) while undergoing both an ultrasound and video recording to directly examine their tongue movements. Participants produced the passage while holding a bite-block between their teeth so that their tongue movements could be viewed from the front, and a camera was placed in front of participants' mouths to take video footage. The passage was designed to contain no labial sounds so as to avoid having the view of the tongue obstructed by the lips. The bite blocks were 5 mm thick and had a small LED light attached to illuminate the oral cavity, as shown on the left of Figure 4. The right side of Figure 4 shows the image captured by the camera.

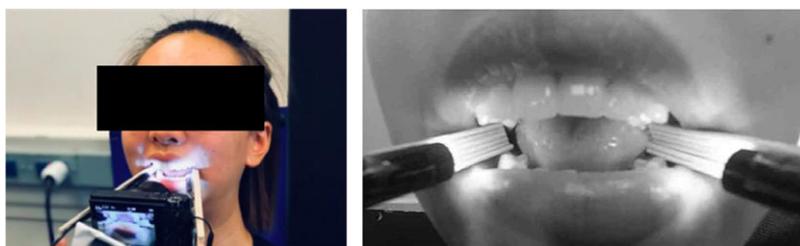


Figure 4 (Colour online) Data collection setup (left) and an image captured by the camera (right) for the validation study.

Each of the validation participants (Val1 and Val2) were seated in a chair with their head resting steady against padded supports. They were asked to hold two bite blocks using their top and bottom premolars. A camera (Sony Cyber-shot DSC RX100) was firmly mounted on an arm attached to the chair and placed in front of the participant's mouth. To help position the camera and the ultrasound probe, participants were asked to repeat 'Mary had a little lamb little lamb little lamb'. The ultrasound probe and camera were placed such that the movements of the sides of the tongue could be clearly viewed through both devices. Both were asked to read the bilabial-free passage (see Appendix) twice in the 5 mm bite block condition. Ultrasound imaging data and camera video footage were collected simultaneously.

For each participant, the repetition that showed the clearest ultrasound and video images most consistently was further analyzed. Video was converted to a grey-scale image sequence and VKGs of the left (shown in the rightmost part on Figure 5), right and center portions of the tongue were produced using ImageJ (Schneider et al. 2012). For the center of the tongue, the VKG slice was taken between the central incisors; for each side of the tongue, the slice was taken at the premolar. A vertical slice was taken from the center of the tongue since that was the direction of movement. But at the sides, the camera showed the tongue was moving towards and away from the teeth at an approximately 45-degree angle, thus a 45-degree angled slice was taken from the sides (shown as black line in the leftmost figure in Figure 5). In each VKG, light areas show periods in which the tongue and teeth (which are lit by the LED) are in contact, while dark areas show an open oral cavity, as shown in the

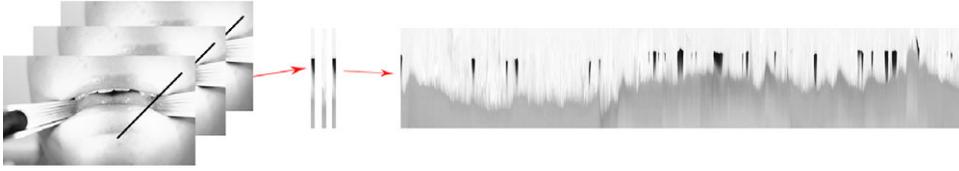


Figure 5 (Colour online) A section of the camera video clip for one side of the tongue (left) and its corresponding VKG (right). The selection includes all visible tongue movement throughout the video clip.

rightmost image in Figure 5. The gray-scale VKG in Figure 5 was then converted to a black and white image by applying a threshold. The same threshold value was applied to VKGs of other regions of the tongue for the same speaker. Hence, black pixels in the black and white VKGs indicate no tongue–palate contact. The total contact duration was determined automatically by counting the number of columns in the VKG without black pixels. The proportion with contact was calculated by dividing the number of columns that had no black pixels by the total number of columns in the VKG. Bracing was determined by comparing the proportion of frames with contact for the sides to that of the center;¹ if the sides make contact while the center does not, then they are providing bracing. A McNemar’s test was performed to test the proportion of frames with right or left tongue–side–palate contact versus tongue–center–palate contact.

Ultrasound imaging data and camera video clips captured simultaneously, ultrasound imaging data was analyzed as described in Section 2.2, and camera video data was analyzed as described in Section 2.3. The ultrasound analysis will be deemed valid if it showed that the sides of the tongue stay at a higher stable region than the center of the tongue during the same periods where the video analysis showed a greater percentage of contact at the sides than the center of the tongue.

3 Results

3.1 Validation

The lingual ultrasound measures for Val1 and Val2 indicated that the median vertical values of the left (Mean = 37.9 mm, SD = 22.36) and right (Mean = 36.8 mm, SD = 23.05) sides of the tongue were smaller than those of the center (Mean = 40.4 mm, SD = 24.86), showing that the sides of the tongue stay in a relatively higher position than the center. Figure 6 shows plots of both the video and ultrasound data. Figure 6a presents the video results showing that the sides of the tongue make contact with the roof of the mouth close to 90 percent of the time, whereas the center of the tongue makes contact with the roof less than 70 percent of the time. A McNemar’s test found the differences of contact duration between the two sides and the center of the tongue were significant for both speakers ($p < .0001$). Figure 6b shows that for the same tokens, the sides of the tongue move in a smaller ISR than the center, which indicates that the sides of the tongue were more stable than the center. Figure 6c shows that the vertical location distributions of the sides of the tongue were high-skewed, whereas the center of the tongue has a low-skewed vertical location distribution, further supporting the view that the sides of the tongue maintain contact with the roof of the mouth.

¹ This measurement over-estimates tongue–palate contact at the center of the tongue since camera view of the pharyngeal wall might be blocked by a high tongue body without touching the palate. However, even the over-estimated medial tongue–roof contact durations are significantly less than the lateral tongue–roof contact, which can be observed in the results section.

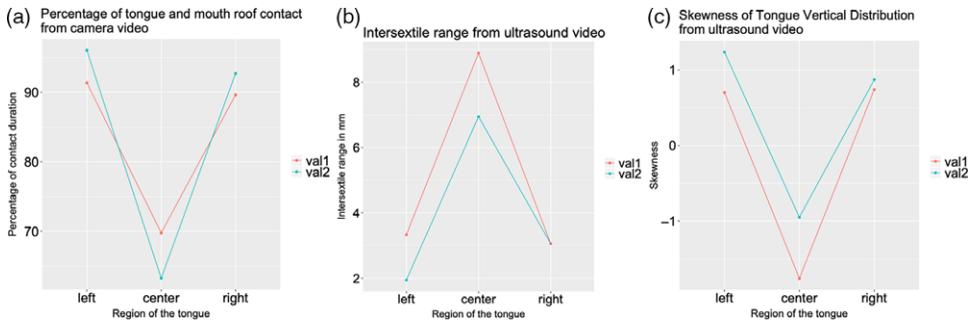


Figure 6 (Colour online) These three figures show results of camera data and ultrasound data from one repetition produced by both validation speakers, represented by Val1 and Val2 respectively in the figures. Figure 6a (leftmost figure) shows the percentage of time that different regions of the tongue contact the roof of the mouth as illustrated by video data and validation analysis. Figure 6b (figure at the center) shows the size of the ISR of different regions of the tongue and Figure 6c (rightmost figure) shows the skewness of the vertical location distribution of different regions of the tongue illustrated by ultrasound data and analysis.

Results from the video analysis shown in Figure 6a thus corroborate the ultrasound results, including the median values as well as the vertical spans and skewness data plotted in Figures 6b and 5c. All of these results are consistent with the view that the lateral tongue maintains nearly constant contact with the upper teeth and palate throughout speech.

This method was used as validation because it showed actual tongue–roof contact, and that lateral bracing was maintained under bite block perturbation for English speakers (Luo et al. 2018). While it was useful for validation, this method was not used for the main cross-linguistic study because it was more invasive, less naturalistic, and difficult to translate a passage into different languages without using labial sounds.

3.2 Size of the ISR

Results of the ISR size of three regions (left, center and right) for each speaker across different languages are shown in Figure 7. ISR represents the range of the center 68% of tongue vertical location distribution. Each chart in Figure 7 shows a different language, and tongue ISR size throughout one repetition. The points for each speaker are joined by separate lines. Wilcoxon Signed rank tests found the ISR size of the center of the tongue was larger than both the left ($p < .0001$, $z = -4.693$) and right side ($p < .0001$, $z = -4.694$) of the tongue. The difference between the left and right sides was also significant ($p < .01$, $z = -2.706$), indicating some lateral asymmetry in tongue release patterns (Liu et al. 2018).

We calculated the mean ISR of the left and right regions (tongue lateral ISR) for each speaker, and compared the effect of languages on tongue lateral ISR using one-way analysis of variance (ANOVA). Average tongue lateral ISR was 9.5 mm for Akan, 12.88 mm for Cantonese, 12.9 mm for English, 15.3 mm for Korean, 12.1 mm for Mandarin, and 10.9 mm for Spanish. ANOVA results suggested that there were no significant differences between tongue lateral ISR of different languages ($F(5) = 0.991$, $p = .446$).

3.3 Median

Whether a region of the tongue stayed higher than others was determined by comparing the medians of vertical position distribution of the region. As a pixel value of zero was assigned to the top of the image, the smaller the pixel value, the higher the vertical position is at. The medians in millimeters of the left side ($M = 21.1$, $SD = 19.9$) and the right side ($M = 21.9$, $SD = 19.8$) were compared to the median of the center of the tongue ($M = 26$, $SD = 19.1$). Wilcoxon Signed Rank tests indicated that the medians of the vertical position distributions

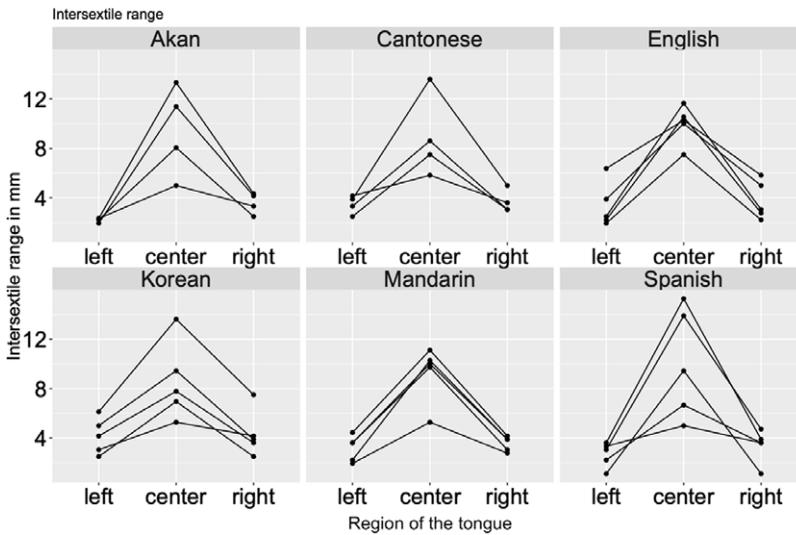


Figure 7 The ISR size of speakers across different languages.

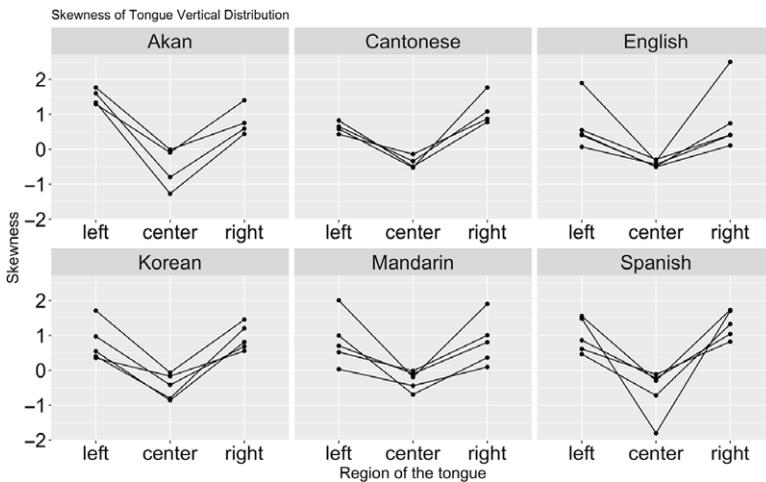


Figure 8 Skewness of the tongue vertical location distributions of all speakers.

of the center of the tongue were significantly larger than the medians of the left ($p < .0001$, $z = -4.431$) and right side ($p = .0003$, $z = -3.587$) of the tongue.

3.4 Skewness

The results of the skewness analysis of the tongue vertical distributions for each speaker are presented in Figure 8. Each chart in Figure 8 shows a different language. Each line represents the skewness of the tongue vertical distribution at the three measurement points on the tongue surface for one speaker throughout one full repetition of the passage. The line graphs show that the skewness of the vertical location distribution of each side of the tongue was greater than 0 for all speakers across different languages, which indicates high-skewed distribution. Results from Shapiro–Wilks test indicate that the vertical location distribution of both sides

of the tongue are high-skewed for all but two speakers. Further, the skewness of the tongue center vertical location distribution for each speaker is at or smaller than 0, indicating a low-skewed distribution. Results from the Shapiro–Wilks test show that distribution of the vertical location of the tongue center was always either normally distributed or low-skewed for each speaker across different languages.

4 Discussion

This study investigates whether lateral tongue bracing is a lingual posture that is pervasively maintained throughout speech in six languages. The results indicate that, for the languages studied, the sides of the tongue tended to stay in a high region most of the time during running speech as reflected by the high-skewed vertical location distribution. The results also illustrate that the sides of the tongue stayed in a higher region than the center of the tongue. Moreover, the smaller sizes of the ISRs of the left and right side of the tongue indicate that they had significantly less vertical movement than the middle, which shows that the sides of the tongue were held comparatively steady at one vertical position in the mouth. Meanwhile, the center of the tongue stays lower and moves more freely, as indicated by the low-skewed vertical location distribution and larger ISR. We interpret these observations as supporting the view that the braced posture of the tongue is persistently maintained during speech in all of the examined languages, consistent with previous findings for English (Bressmann et al. 2010, Gick et al. 2017), as well as other languages (Cheng et al. 2017). Moreover, our results suggest that lateral bracing is maintained pervasively in these languages rather than only occurring in certain sounds (e.g. Kwok & Stokes 1997, Kim 2001, Kochetov & Colantoni 2011, Yu et al. 2013). These findings support the view that lateral bracing is a general property of speech production rather than an aspect of one particular language.

Both ISR size and skewness results show variation in all three regions among speakers of the same language. Aside from some expected random inter-speaker variation, there are additional reasons that might contribute to such variation. For instance, the target tongue area was not exactly the same for different speakers during data collection, which would contribute to variation in both ISR size and skewness. In addition, the number of data points remaining after applying thresholds to the ultrasound VKG varied across speakers. However, as the same threshold was applied to all three tongue regions, the within-speaker variation for these measures is believed to be small. Though females typically have a smaller tongue than males (Stone 2005), no significant difference was observed between ISR size of female (5.49) and male (5.35) speakers ($p = .82$).

For the Sun and Wind passages, Cantonese had the largest proportion (22%) of laterals and low vowels (as shown in Table 1), followed by Spanish (21%), Mandarin (19%), Korean (18%), Akan (16%) and English (12%). As release of bracing was observed for laterals and low vowels (Gick et al. 2017) and the proportion of these sounds varies across languages, the amount of lateral tongue vertical movement might be expected to differ across languages. Yet, no significant difference in ISR between any of the languages was found, which might be due to the small number of speakers for each language. Also, in Gick et al. (2017), release only occurred for a quarter (25.2%) of laterals and 18.2% of low vowels (/ɑ aʊ ai ʌ/) in English, and varied between the speakers. Thus, the proportion of released laterals and low vowels is expected to be different across speakers and languages. In the present study, the proportion released could not be determined since the upper teeth were not visible in the ultrasound imaging, thus whether laterals and low vowels were produced in a braced or released manner could not be determined.

If lateral bracing predominates throughout continuous speech across languages, that would suggest that bracing is a physiological property of speech production, potentially shedding light on speech research more broadly, with relevance to clinical practice (e.g. Bernhardt et al. 2005). On the one hand, McLeod (2011) found that a sample of Speech

Language Pathologists (SLPs) showed limited awareness of lateral bracing for alveolar consonants including /t d n s z/ and posterior lateral contact for other consonants including /b v θ ð w/. On the other hand, the SLPs showed better knowledge of posterior lateral contact for /j/ and /ɹ/ (McLeod 2011). It is important for clinicians to have knowledge of bracing because accurate tongue–palate patterns have been considered indicative of successful treatment, for example, for the misarticulation of /s/ (McAuliffe & Cornwell 2008).

Our findings also shed light on modeling of speech and other oral behaviors, suggesting that lateral bracing should be taken into account in biomechanical simulations of speech; specifically, lateral bracing should be maintained for most sounds during simulations of speech production. Likewise, Mayer et al. (2017) ran simulations showing that lateral bracing for speech is controlled in very similar ways to bracing postures associated with feeding and swallowing, contributing to the notion that lateral tongue bracing may be a deeply rooted physiological property underlying both feeding/swallowing and speech behaviors.

This study has potential limitations. Though this study investigates a variety of different languages, the number of speakers and the range of languages is limited, and future work looking at additional languages from more speakers may be fruitful. Moreover, due to the nature of ultrasound imaging, tongue contour was not always visible in the regions of interest, and so some position data was lost. Additionally, the ultrasound probe was placed at somewhat different angles from speaker to speaker, which might account for some of the inter-speaker variation. Finally, though participants' heads were stabilized, some head movement was possible and may have resulted in some vertical shifts in tongue position.

5 Conclusion

This study investigated lateral tongue movement using ultrasound imaging during running speech for six different languages. It introduced and validated a new methodology that is significantly easier and more cost-effective than other ways of measuring bracing such as EPG. The methodology developed for this study could also be used to investigate movement patterns of visible objects captured by camera, ultrasound or Structural MRI. Specifically, it could track back-and-forth movements of a single visible object. For example, the VKG could track the movement of the velum in Structural MRI and measure the extent of velopharyngeal port opening. The results show that the sides of the tongue remained within a limited higher region whereas the center of the tongue stayed relatively lower and made larger movements than the sides for all speakers. These findings indicate that the sides of the tongue were braced against the sides of the roof of the mouth much of the time across the different languages, supporting the view that lateral tongue bracing is a fundamental feature of human speech production. Further studies could investigate whether lateral bracing is maintained symmetrically at both sides and the mechanism required for releasing and raising the sides during a disruption of lateral bracing posture.

Acknowledgements

This research was funded by the National Institute of Health (NIH) Grant DC-002717, and National Science and Engineering Research Council of Canada (NSERC) RGPIN-2015-05099 and RGPIN-2021-03751. We gratefully acknowledge the comments made by the associate editor, Dr Alexei Kochetov, and three anonymous reviewers. We also thank Dawoon Choi for the helpful discussions.

Appendix. Passage used in the bite block experiment

Today is a nice day. I decided to take a taxi to the city to see Alex. Alex is a dentist. He has six dogs.

I saw Alex at a diner and had tea and cake. Then, Alex and I headed to the clinic. He said that there's a kid, Dan, that hates the dentist. He gets sad and yells a lot and doesn't let Alex clean his teeth.

Today, Dan is at the clinic again.

'I hate the dentist!' Dan yells at Alex, and he still doesn't let Alex clean his teeth. Suddenly, Alex's dog, Dex, sits next to Dan and licks his hand. Dan giggles and says to Alex, 'Dex has nice teeth! I need teeth like that!' Then, he lets Alex clean his teeth. Alex is stunned.

References

- Amoako, Wendy Kwakye. 2020. Assessing phonological development among Akan-speaking children. Master's thesis, University of British Columbia.
- Avelino, Heriberto. 2018. Mexico City Spanish. *Journal of the International Phonetic Association* 48(2), 223–230.
- Bernhardt, Barbara, Bryan Gick, Penelope Bacsfalvi & Marcy Adler-Bock. 2005. Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics & Phonetics* 19(6–7), 605–617.
- Bressmann, Tim, Heather Flowers, Willy Wong & Johnson C. Irish. 2010. Coronal view ultrasound imaging of movement in different segments of the tongue during paced recital: Findings from four normal speakers and a speaker with partial glossectomy. *Clinical Linguistics & Phonetics* 24(8), 589–601.
- Cheng, Hei Yan, Bruce E. Murdoch, Justine V. Goozee & Dion Scott. 2007. Electropalatographic assessment of tongue-to-palate contact patterns and variability in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research* 50(2), 375–392.
- Cheng, Laurreta, Murray Schellenberg & Bryan Gick. 2017. Cross-linguistic bracing: A lingual ultrasound study of six languages. *Canadian Acoustics* 45(3), 186–187.
- Corneau, Caroline. 2000. An EPG study of palatalization in French: Cross-dialect and inter-subject variation. *Language Variation and Change* 12(1), 25–49.
- D'Ausilio, Alessandro, Eleonora Bartoli, Laura Maffongelli, Jeffrey James Berry & Luciano Fadiga. 2014. Vision of tongue movements bias auditory speech perception. *Neuropsychologia* 63, 85–91.
- Dolphyne, Florence Abena. 1988. *The Akan (Twi-Fante) language: Its sound systems and tonal structure*. Accra: Ghana Universities Press.
- Dryer, Matthew S. & Martin Haspelmath. 2013. The world atlas of language structures online. Retrieved from <https://wals.info/>.
- Gibbon, Fiona E., Ivan Yuen, Alice Lee & Lynne Adams. 2007. Normal adult speakers' tongue palate contact patterns for alveolar oral and nasal stops. *Advances in Speech Language Pathology* 9(1), 82–89.
- Gick, Bryan, Blake Allen, François Roewer-Després & Ian Stavness. 2017. Speaking tongues are actively braced. *Journal of Speech, Language, and Hearing Research* 60(3), 494–506.
- Gick, Bryan, Ian Wilson & Donald Derrick. 2013. *Articulatory phonetics*. Malden, MA: John Wiley & Sons.
- Ha, Seunghee, Cynthia J. Johnson & David P. Kuehn. 2009. Characteristics of Korean phonology: Review, tutorial, and case studies of Korean children speaking English. *Journal of Communication Disorders* 42(3), 163–179.
- Hillenbrand, James M. 2003. American English: Southern Michigan. *Journal of the International Phonetic Association* 33(1), 121–126.
- IPA [International Phonetic Association] (ed.). 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Kim, Dae Won. 2000. Context-sensitive lingual gestures in the Korean tap /ɾ/. *Speech Sciences* 7(3), 11–19.
- Kim, Hyunsoon. 2001. The place of articulation of the Korean plain affricate in intervocalic position: An articulatory and acoustic study. *Journal of the International Phonetic Association* 31(2), 229–257.
- Kochetov, Alexei & Laura Colantoni. 2011. Coronal place contrasts in Argentine and Cuban Spanish: An electropalatographic study. *Journal of the International Phonetic Association* 41(3), 313–342.

- Kwok, Irene C. L. & Stephanie F. Stokes. 1997. An electropalatographic description of Cantonese speech. *Asia Pacific Journal of Speech, Language and Hearing* 2(2), 125–138.
- La'Porte, Sarah J., Jaspal K. Juttla & Ravi K. Lingam. 2011. Imaging the floor of the mouth and the sublingual space. *Radiographics* 31(5), 1215–1230.
- Lee, Hyun Bok. 1993. Korean. *Journal of the International Phonetic Association* 23(1), 28–31.
- Lee, Wai Sum & Eric Zee. 2003. Standard Chinese (Beijing). *Journal of the International Phonetic Association* 33(1), 109–112.
- Liu, Yadong, Megan Keough, Felicia Tong, Oksana Tkachman, Kate Radford & Bryan Gick. 2018. Lateral bias in lingual bracing during speech. *Canadian Acoustics* 46(4), 56–57.
- Lulich, Steven M. & William G. Pearson. 2019. Three-/four-dimensional ultrasound technology in speech research. *Perspectives of the ASHA Special Interest Groups* 4(4), 733–747.
- Lundberg, Andrew J. & Maureen Stone. 1999. Three-dimensional tongue surface reconstruction: Practical considerations for ultrasound data. *The Journal of the Acoustical Society of America* 106, 2858–2867.
- Luo, Sophia, Yadong Liu, Arian Shamei, Murray Schellenberg, Monika Łuszczuk & Bryan Gick. 2018. Tongue bracing under bite block perturbation. *Canadian Acoustics* 46(4), 54–55.
- Mayer, Connor, François Roewer-Després, Ian Stavness & Bryan Gick. 2017. Does swallowing bootstrap speech learning? *Canadian Acoustics* 44(3), 194–195.
- McAuliffe, Megan J. & Petrea L. Cornwell. 2008. Intervention for lateral /s/ using electropalatography (EPG) biofeedback and an intensive motor learning approach: A case report. *International Journal of Language & Communication Disorders* 43(2), 219–229.
- McLeod, Sharynne. 2006. Australian adults' production of /n/: An EPG investigation. *Clinical Linguistics & Phonetics* 20(2–3), 99–107.
- McLeod, Sharynne. 2011. Speech-language pathologists' knowledge of tongue/palate contact for consonants. *Clinical Linguistics & Phonetics* 25(11–12), 1004–1013.
- McLeod, Sharynne, Amber Roberts & Jodi Sita. 2006. Tongue/palate contact for the production of /s/ and /z/. *Clinical Linguistics & Phonetics* 20(1), 51–66.
- Pulleyblank, Douglas. 2018. Niger-Kordofanian (Niger-Congo) languages. In Bernard Comrie (ed.), *The world's major languages*, 3rd edn., 877–924. Oxford: Routledge.
- R Core Team. 2019. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Schneider, Caroline A., Wayne S. Rasband & Kevin W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9(7), 671–675.
- Song, Jae Jung. 2006. *The Korean language: Structure, use and context*. Abingdon: Routledge.
- Stevens, Kenneth N. & Joseph S. Perkell. 1977. Speech physiology and phonetic features. In Masayuki Sawashima & Franklin S. Cooper (eds.), *Dynamic aspects of speech production: Current results, emerging problems, and new instrumentation*, 323–341. New York: Columbia University Press.
- Stone, Maureen. 1990. A three-dimensional model of tongue movement based on ultrasound and X-ray microbeam data. *The Journal of the Acoustical Society of America* 87(5), 2207–2217.
- Stone, Maureen. 2005. A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics & Phonetics* 19(6–7), 455–501.
- Švec, Jan G. & Harm K. Schutte. 1996. Videokymography: High-speed line scanning of vocal fold vibration. *Journal of Voice* 10(2), 201–205.
- Yu, Hongzhi, Yanhua Duan & Dongxu Zhang. 2013. The research on Mandarin Chinese consonants from perspective of EPG. *International Workshop on Computer Science in Sports*, 32–35.
- Zee, Eric. 1991. Chinese (Hong Kong Cantonese). *Journal of the International Phonetic Association* 21(1), 46–48.
- Zsiga, Elizabeth C. 1995. An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In Bruce Connell & Amalia Arvaniti (eds.), *Phonology and phonetic evidence* (Papers in Laboratory Phonology IV), 282–302. Cambridge: Cambridge University Press.