Speech Communication 113 (2019) 15–24

Contents lists available at ScienceDirect

Speech Communication

journal homepage: www.elsevier.com/locate/specom

# Differentiating tongue shapes for alveolar-postalveolar and alveolar-velar contrasts

Natalia Zharkova

*Speech and Language Sciences, Newcastle University, King George VI Building, Newcastle upon Tyne NE1 7RU, United Kingdom*

ARTICLE INFO

*Keywords:*
Tongue shape
Ultrasound
Alveolar
Postalveolar
Velar

ABSTRACT

This paper is focussed on differentiating midsagittal tongue shapes for alveolar-postalveolar and alveolar-velar contrasts in place of articulation. In addition to two established measures assessing the shape of a tongue curve, three new indices are introduced, which capture more fine grained distinctions in tongue shape, through quantifying the extent of curvature at different locations along the tongue contour. In order to establish whether the indices can be applied across a range of ultrasound recording settings, including those where the transducer is hand-held rather than stabilised in relation to the head, the length of the tongue curve was systematically manipulated, since differences in the length of the imaged tongue contour are likely to occur in non-head-stabilised recordings. The study identified the measures that differentiated between contrasting places of articulation, producing robust results regardless of differences in the length of the imaged tongue curve. Combinations of measures were found to capture place of articulation distinctions more accurately than individual measures. The paper includes a discussion of how such combinations of measures can be used for studying the development of phonological contrasts in typical child speech, as well as the realisation of alveolar-postalveolar and alveolar-velar oppositions in disordered speech.

## 1. Introduction

Contrasts between lingual places of articulation are of particular interest in studying both typical and disordered speech, especially speech produced by young children. For example, velar fronting, i.e., producing /k/ as [t], and postalveolar fricative fronting, i.e., producing /ʃ/ as [s], together account for many errors in the speech of young typically developing English speaking children (e.g., Smit, 1993), as well as for persistent errors in children with speech sound disorders (Dodd, 2013). Another common phonological process in both typical and disordered speech development in English consists in producing target liquid consonants /l/ and /r/ as a labial-velar approximant, which involves a different tongue shape from that of the liquid target, in addition to the labial articulation. Lingual place of articulation substitutions are common cross-linguistically, although the direction of the most typical substitutions in speech acquisition may be the opposite to the fronting processes described above, as with the backing of stops from dental to velar, or the backing of fricatives from alveolar to postalveolar in Japanese (see, e.g., Beckman et al., 2003). An accurate characterisation of tongue shape during perceived substitutions can help establish whether a perceived phonological substitution might in fact be a covert contrast (see, e.g., Macken and Barton, 1980). Covert contrasts involve sounds that are produced by the speaker with a measurable difference, but that are “heard and transcribed by listeners with the same phonetic symbol” (Gibbon and Lee, 2017a: 1). Identifying covert contrasts can lead to a more accurate characterisation of the production of interest, and potentially to an earlier acquisition of the contrast in delayed or disordered development (Tyler et al., 1993).

An increasing number of studies have used acoustic analysis and/or electropalatography to trace covert contrasts in place of articulation in young children’s productions (e.g., Hewlett, 1988; Edwards et al., 1997; Li et al., 2009; Munson et al., 2010; Gibbon and Lee, 2017b, and references cited there). Ultrasound tongue imaging is potentially also a very suitable technique, since it can directly show differences in tongue shape across speech sounds. In a small ultrasound study focussed on /k/ and /t/ productions by 4-year-old children, McAllister Byun et al. (2016) demonstrated some evidence of covert contrasts in one of two children who exhibited the phonological process of velar fronting. The authors used Dorsum Excursion Index (DEI), a measure developed by Zharkova (2013). DEI quantifies the extent of midsagittal tongue bunching, and Zharkova (2013) showed that the velar stop /k/ had greater DEI values than several alveolar consonants in the context of /a/. This measure, together with some other tongue shape measures, was identified by Zharkova et al. (2017) as potentially suitable for studying the development of phonological oppositions and identifying covert contrasts in young children’s speech. In the current study, several tongue shape measures were applied to analysing two linguistically relevant

*E-mail address:* natalia.zharkova@newcastle.ac.uk

https://doi.org/10.1016/j.specom.2019.08.004
Received 2 January 2019; Received in revised form 18 July 2019; Accepted 12 August 2019
Available online 13 August 2019

contrasts in tongue shape: alveolar-velar and alveolar-postalveolar contrasts in English. In addition to DEI, which was used to assess alveolar-velar distinctions, an index called $LOC_{a\text{-}i}$ (Zharkova et al., 2015) was employed, since it had previously been shown to differentiate between /s/ and /ʃ/ tongue shapes in children and adults (e.g., Zharkova, 2016). While DEI provides information on the overall extent of bunching from a midsagittal tongue contour, $LOC_{a\text{-}i}$ assesses how much tongue bunching occurs towards the front half of the tongue contour, compared with the back half. This index has previously been used to differentiate between midsagittal tongue shapes for alveolar and postalveolar fricatives in adults' and preadolescent children's speech, with /ʃ/ reported to have higher $LOC_{a\text{-}i}$ values than /s/ (Zharkova, 2016). A more recent study by Zharkova et al. (2018) extended those findings to a wider range of child ages. Both DEI and $LOC_{a\text{-}i}$ are described in detail below in Section 2.3.1, including illustrations in Fig. 4.

DEI and $LOC_{a\text{-}i}$ are based on ratios of straight lines characterising the shape of a midsagittal tongue curve (cf. Bressmann et al., 2005; Aubin and Ménard, 2006). Three new measures were additionally designed in this study, with the aim to capture more local aspects of the tongue curve that were expected to differ across the target consonants. These expectations were driven by previously reported findings on tongue movements for those consonants (e.g., Zharkova, 2007; 2013; 2016; Recasens and Rodríguez, 2016; 2018b). Unlike $LOC_{a\text{-}i}$ and DEI, the new measures are based on quantifying curvature at specific locations along the front two thirds of the tongue curve, using three selected points on the tongue contour to define those locations. The new indices are described in detail in Section 2.3.2.

All the measures in the present study capture aspects of tongue shape based on information from a single tongue curve. For statistical analysis purposes, this means that individual tongue curves from multiple repetitions of a target or multiple targets (e.g., several repetitions of [t] from "tea" and several repetitions of [k] from "key") do not need to be located within the same coordinate space for comparing tongue shapes across targets. This, in turn, makes it possible to have some flexibility in the process of data collection. In particular, there is no requirement that the ultrasound transducer is stabilised in relation to the speaker's head, or that some arrangements are made at the recording stage to enable subsequent head-to-transducer correction. The single curve based approach thus has an advantage over a number of other measurement approaches, such as whole curve based Smoothing Spline ANOVA (Davidson, 2006) or Nearest Neighbour distances (Zharkova and Hewlett, 2009), and measuring horizontal coordinates of the "highest point" on the tongue curve (e.g., Ménard et al., 2012; Noiray et al., 2018; Abakarova et al., 2018; Rubertus and Noiray, 2018), all of which require that multiple curves are located in the same coordinate space. Since indices based on characterising the shape of a single tongue curve are potentially less powerful in capturing linguistically relevant differences than measures based on comparing whole tongue curves in the same coordinate space (see a discussion in Zharkova et al., 2015), the three additional measures were expected to provide complementary information on fine differences in tongue shape between contrasting places of articulation.

The application of tongue shape measures to the speech of young children requires the measures to provide robust results regardless of differences in absolute positions of tongue contours across repetitions and across target speech sounds (see, e.g., Zharkova et al., 2015). Measures based on a single tongue contour have previously been shown to provide reliable results regardless of tongue curve orientation and location in a coordinate space (e.g., Ménard et al., 2012; Zharkova et al., 2015; see also a recent study comparing single-curve-based measurements across ultrasound and MRI data, Kansy et al., 2018). However, there have been no studies, to date, addressing the question whether such measures would be affected by differences in imaged tongue length. Such differences constitute a plausible outcome of ultrasound recordings of small children without head-to-transducer stabilisation (see, e.g., Zharkova et al., 2017). Therefore the current study analysed the performance of each of the measures across simulated conditions differing in the imaged tongue length. For each of the two lingual contrasts of interest, the study aimed to identify the measures that can robustly differentiate between contrasting tongue shapes, regardless of the curve length differences.

## 2. Materials and methods

### 2.1. Datasets

All the data were from adult speakers of Scottish Standard English, without speech disorders. The alveolar-postalveolar contrast was analysed using a different dataset from that used for the alveolar-velar contrast, and the two datasets are described in detail below. The speakers from each dataset were selected to minimise across-speaker differences in age and gender distribution between the two datasets. Since no comparisons were carried out across the two datasets, the fact that the data for the two lingual contrasts were from different groups of speakers was not a problem for the analysis, but this difference needs to be borne in mind when visually inspecting the tongue curves from the two datasets. In each dataset, the data from five repetitions of each target consonant were included in the analyses, with tongue curves traced at mid-consonant (for the stops, at mid-closure) used for calculating the tongue shape indices. Both datasets were collected using a head-to-transducer stabilising headset designed by Articulate Instruments Ltd.

In order to compare tongue shapes in alveolar and postalveolar fricatives, a subset of the data was selected from the corpus described in Zharkova (2016). In this subset, six participants were selected, four female and two male. The mean age was 35 years old, and the range was between 22 and 46 years old. The speakers produced each of the target fricatives, /s/ and /ʃ/, in CV syllables, with two different vowels, /a/ and /i/, embedded in a carrier phrase. The /s/-/ʃ/ consonant pair was analysed in each of the two vowel contexts. The dataset including /s/ and /ʃ/ is henceforth referred to as the alveolar-postalveolar dataset. The tongue contours for this dataset are plotted in Figs. 1 and 2, in the context of /a/ and in the context of /i/, respectively. The tongue contours for both datasets were traced semi-automatically, with manual correction, using Articulate Assistant Advanced software (Articulate Instruments Ltd, 2012). The front and back ends of the tongue curve were determined by locating the intersection of the lower edge of the bright white line (corresponding to the tongue surface) in the ultrasound image and the shadows of the chin and hyoid bones, respectively.

The data for the velar-alveolar contrast were taken from a corpus described in Zharkova (2013). The participants were six speakers, with the same gender distribution and mean age as in the alveolar-postalveolar dataset, and the age range between 23 and 46 years old. The target consonants in this study were /k/, /t/, /l/ and /r/[1], produced at the onset of short words ("cast", "tasked", "lasted" and "raps", respectively; the vowel immediately following the target consonant was always /a/), embedded in carrier phrases. The context of /i/, also present in the original corpus, is not reported in this study, because the contrast between velar and alveolar tongue shapes was much less noticeable in the high vowel context, due to the tongue predorsum raising for the alveolars, and consequently none of the indices showed a robust difference between tongue shapes of the velar stop and the alveolar consonants. The dataset including /k/ and the alveolar consonants in the context of /a/ is henceforth referred to as the alveolar-velar dataset. The tongue contours for this dataset are plotted in Fig. 3.

[1] Note that the indices described in this paper aim to capture a distinction between alveolar and velar places of articulation in /a/ context, without directly addressing the question whether the shape of the tongue is more or less complex (cf. previous studies that have used quantitative indices designed to focus on measuring tongue shape complexity, particularly as applied to rhotic phonemes: Gick et al., 2008; Dawson et al., 2016; Preston et al., 2019).

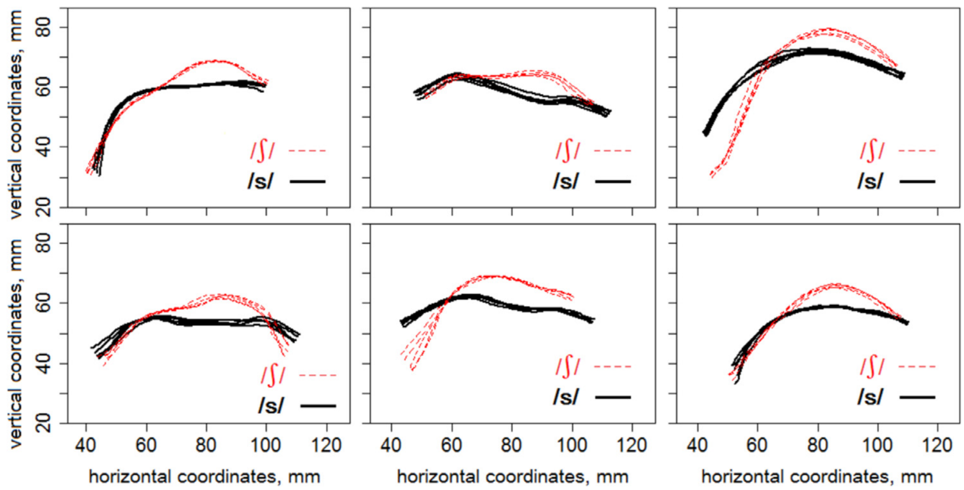


**Fig. 1.** Tongue contours for five repetitions of /s/ (thick solid curves) and five repetitions of /ʃ/ (thin dashed curves) in the context of /a/. Each graph corresponds to one speaker from the alveolar-postalveolar dataset. The front of the tongue is on the right in this figure, as well as in Figs. 2–6.

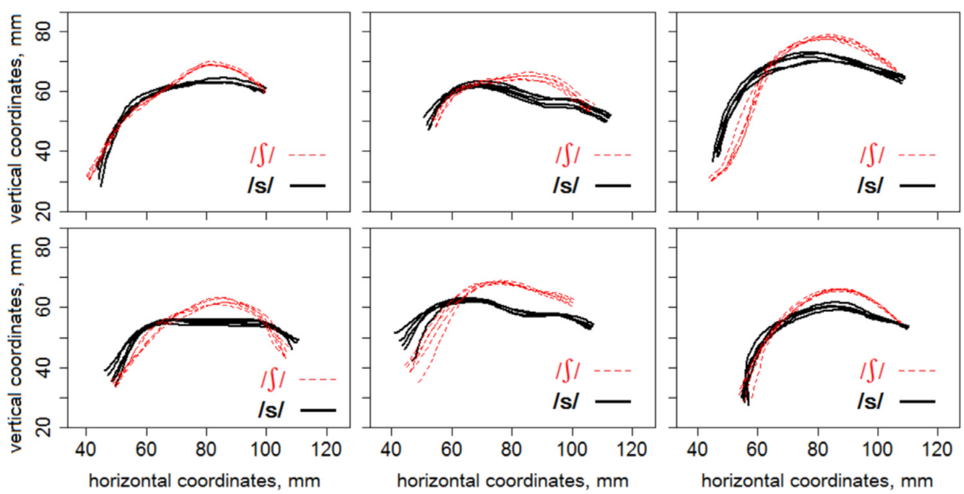


**Fig. 2.** Tongue contours for five repetitions of /s/ (thick solid curves) and five repetitions of /ʃ/ (thin dashed curves) in the context of /i/. Each graph corresponds to one speaker from the alveolar-postalveolar dataset. The order of speakers represented in the six panels is the same as in Fig. 1.

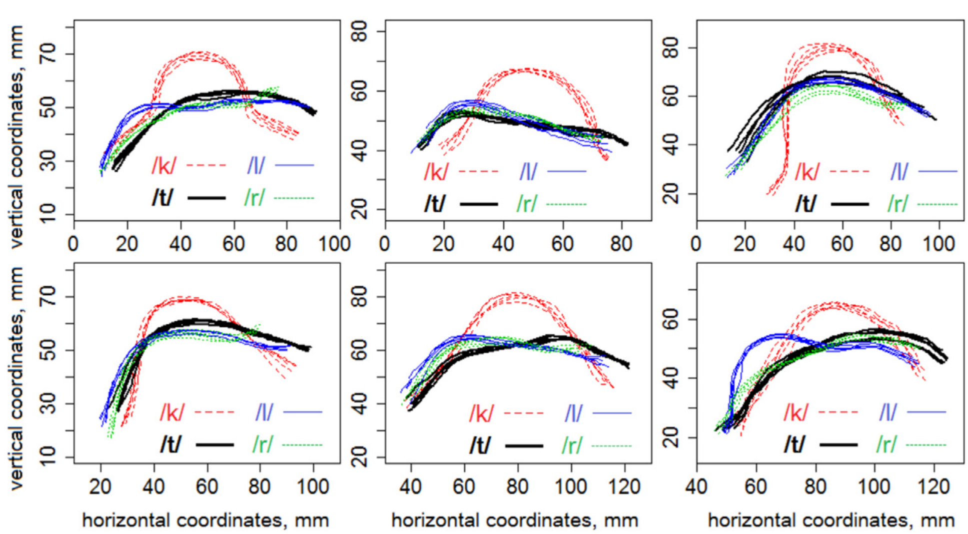


**Fig. 3.** Tongue curves for /k/ (thin dashed curves), /t/ (thick solid curves), /l/ (thin solid curves) and /r/ (thin dotted curves) in the context of /a/, for each speaker from the alveolar-velar dataset. The scales are different across speakers in order to facilitate visualisation.

### 2.2. Manipulating curve length

For assessing whether across-consonant differences in tongue shape were affected by differences in the imaged curve length, the same tongue shape indices were calculated in four different conditions. The first condition involved the whole tongue curve (thereafter referred to as “Full”). In the other three conditions, a part of the tongue curve was missing. One of the challenges in ultrasound analysis of tongue movements relates to imaging the tongue tip, and it has been established that up to one centimetre of the tip may not be imaged due to air underneath the tongue tip (Stone, 2010). Therefore in the second condition in the present study a centimetre at the front of the curve

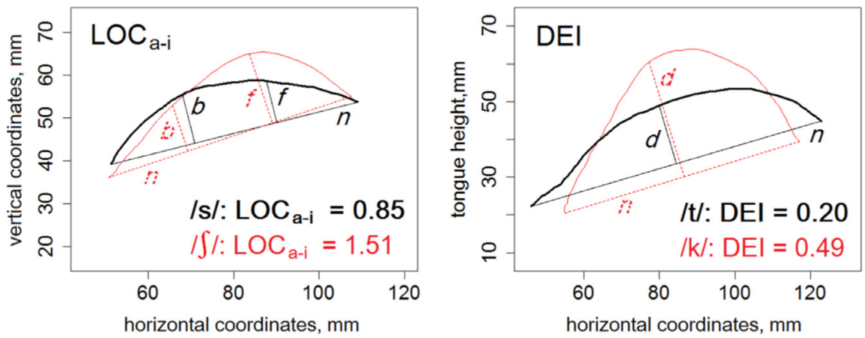


**Fig. 4.** Example tongue curves, illustrating the calculation of LOC$_{a\text{-}i}$ (left panel) and DEI (right panel). In the left panel, a solid thick curve is for /s/ from /sa/, with superimposed solid straight lines that are used for computing the index for the alveolar fricative (see further explanations in the text); a solid thin curve is for /ʃ/ from /ʃa/, with superimposed dashed straight lines used for computing the index for the postalveolar fricative. In the right panel, a solid thick curve with superimposed solid straight lines is for /t/ from /ta/, and a solid thin curve with its associated dashed straight lines is for /k/ from /ka/.

was missing (thereafter referred to as "Tm", for "Tip missing"). In the third condition, symmetrically to the second one, a centimetre was missing at the other end of the curve, to represent a possibility of the shadow of the hyoid bone obscuring more of the tongue curve above it (referred to as "Bm", for "Back missing"). Finally, a centimetre was missing at the front of the curve and another centimetre was missing at the back of the curve, resulting in the total of two centimetres of the curve excluded from the data (referred to as "T&Bm", for "Tip and Back missing"). The tongue curves for those three conditions were obtained by shortening the original curves from the Full condition, using the scripts written in R (R Core Team, 2013) by the author.

### *2.3. Articulatory indices*

#### *2.3.1. LOC$_{a\text{-}i}$ and DEI*

LOC$_{a\text{-}i}$ captures the difference in the extent of bunching between the front and the back of the tongue, with higher values representing tongue shapes with more excursion of the front than the back of the tongue (for example, for a close front vowel). DEI quantifies the extent of bunching of mid-tongue, with higher values corresponding to more bunched tongue shapes (such as that for a velar consonant), and lower values corresponding to flatter tongue shapes (such as that for an alveolar consonant in the context of an open vowel). Fig. 4 illustrates the calculation of the two indices. LOC$_{a\text{-}i}$ is the ratio of *f* to *b* (see the left panel of the figure). The two lines, *f* and *b*, are perpendiculars to the tongue curve from two points located on line *n* (a straight line between the two ends of the tongue curve), at one third and two thirds from the front end of the curve, respectively. DEI is the ratio of *d* to *n* (see the right panel), where *d* is a perpendicular to the tongue curve from mid-*n*.

#### *2.3.2. TF1, TF2 and TF3*

In addition to LOC$_{a\text{-}i}$ and DEI, three new indices were developed for the purpose of capturing fine differences in tongue shape across contrasting lingual places of articulation.

Examining Figs. 1–3, we can see that in the front half to two thirds of the tongue there are particularly noticeable differences, across speakers, between /s/ and /ʃ/, as well as between /k/ and alveolar consonants. This study aimed to differentiate between those contrasting tongue shapes by sampling the front two thirds of the tongue contour more densely. For all three new indices, a somewhat different approach was followed to that employed in calculating LOC$_{a\text{-}i}$ and DEI. Instead of tracing a straight line between two ends of the curve, three points were chosen along the tongue surface for calculating each index (Fig. 5 shows example tongue curves with the three points connected by straight lines). These three points were used to find the length of the radius of a circle passing through these points, and to identify the extent of concavity of the circumference of this circle. In each case, the inverse of the radius of the circle passing through the three points was calculated, and the resulting value was multiplied by ten (for the practical reason of reducing the number of decimal places), yielding the index value. In Fig. 6, an enlarged section of the top left panel from Fig. 5 contains further annotations, with the points making up the triangles referred to as points A, B and C. The angle at point C can be used for visualisation purposes, with larger values of this angle corresponding to a circle of a greater circumference, and, consequently, to a smaller value of the index (in Fig. 6, the index TF1 is used as an example, but the same principle applies to calculating all three new indices). This approach is similar to that used by Stone et al. (1987) to quantify aspects of tongue shape. An important difference is that Stone and colleagues computed curvature for multiple sets of three points for each curve, and used the resulting numbers in combination with absolute data on the tongue

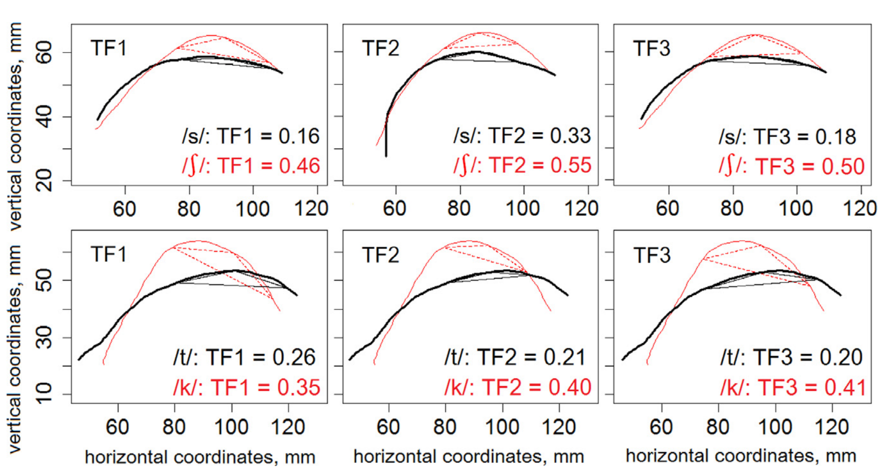


**Fig. 5.** Example tongue curves, illustrating the calculation of TF1 (left), TF2 (centre) and TF3 (right). The top row: a thick curve for /s/, and a thin curve for /ʃ/ (the context of /a/ in the left and right panels, and the context of /i/ in the central panel); the bottom row: a thick curve for /t/ from /ta/, and a thin curve for /k/ from /ka/. In all six panels, the triangle made of solid lines represents the calculations for the alveolar consonant, and the triangle made of dashed lines represents the calculations for the other consonant.

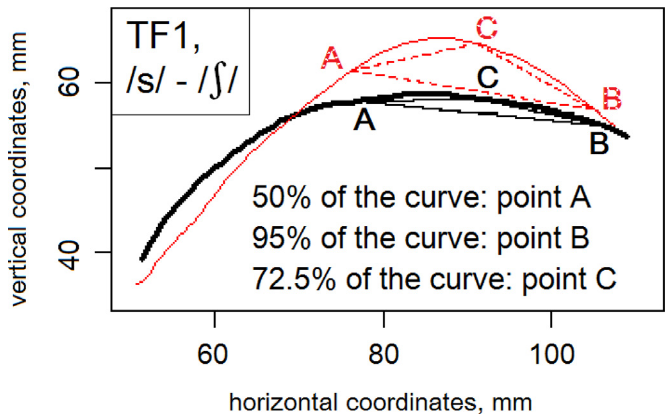


**Fig. 6.** Annotated example tongue curves illustrating the calculation of TF1 (the curves are the same as those presented in the top left panel of Fig. 5: a thick curve for /s/ from /sa/, and a thin curve for /ʃ/ from /ʃa/); see the text for details.

position. In the present study, the aim was to compare curve shapes without reference to their absolute position, so as to be able to use the indices for non-head-stabilised data, and for comparing tongue shapes across speakers. Selecting three points along the curve for each index makes it possible to quantify aspects of tongue shape in specific parts of the tongue across speech segments and participants. The indices, described in detail below, were calculated in R (R Core Team, 2013), using the scripts written by the author.

The calculations of the index referred to as TF1 ("TF" stands for "tongue front") are illustrated in the left panels of Fig. 5. For establishing the location of each of the three points for this index, as well as the points used for calculating the other radius-based indices, the length of the curve was first calculated, by summing up Euclidian distances between each successive pair of xy coordinates. For calculating TF1, the xy coordinates were then identified for the points located at mid-curve and at 95% of the curve, starting from the back (henceforth, in the description of the radius-based indices, the locations of the points refer to the distance starting from the back of the curve, so higher numbers represent locations further forward along the curve). The third, middle, point was located half-way between the other two points, at 72.5% of the curve. Since the part of the circle passing through these three points spanned the region between the blade[2] and the dorsum of the tongue, the index was expected to capture the difference between /s/ and /ʃ/ in the shape of the front of the tongue. Specifically, it was hypothesised that higher values of this index would be observed for /ʃ/ than for /s/, resulting likely from the lowering and retraction of the tip of the tongue for /ʃ/ in order to create a sublingual cavity, as opposed to a flatter midsagittal shape of the front half of the tongue for /s/. This pattern is indeed illustrated in the top left panel of Fig. 5. Further details on calculating TF1 for this consonant contrast are available in Fig. 6, which shows tongue curves for /s/ from /sa/ and /ʃ/ from /ʃa/. The angle at the middle point of the triangle (point C in Fig. 6) is noticeably larger for /s/, and TF1 value is, accordingly, smaller for this consonant than for the postalveolar fricative.

The calculations of TF2 and TF3 are illustrated in the central panels and the right panels of Fig. 5, respectively. Both these indices had their middle point centred further back along the tongue curve than for TF1, namely at two thirds of the curve. For TF2, the other two points were located at mid-curve (i.e., further back than the point at two thirds of the curve) and at the same distance in the forward direction, i.e., at 83.4% of the curve. Differences in this region of the tongue curve could be expected to occur between alveolars and consonants articulated further back in the oral cavity, with flatter shapes for the alveolars. In the two central panels in Fig. 5, alveolar consonants show consistently different shapes from those of the consonants articulated further back in the vocal tract. This is represented in TF2 values, and the difference can be visualised by comparing the angle at the middle point of the triangle across consonants, in each of the two panels. For TF3, also centred at two thirds of the curve, the other two points were located further away from the middle point, specifically, at equal distances from it, at 43.4% and 90% of the curve length. The part of the imaginary circle passing through the three points encompassed a greater proportion of the tongue curve than that for TF2, and TF3 index was expected to capture relatively more global differences in the front of the tongue shape between alveolars and non-alveolars, differentiating between tongue shapes with more versus less bunching of the tongue predorsum. This pattern is represented in the index values shown in the two panels on the right in Fig. 5, with alveolar consonants having larger angles at the middle point of the triangle than non-alveolar consonants, and, consequently, smaller TF3 values.

### 2.4. Statistical analyses

All statistical analyses were carried out in R (R Core Team, 2013). The analyses were run separately for the two datasets, and, in the case of the alveolar-postalveolar dataset, separately for the two vowel contexts.

In order to evaluate the performance of the indices in distinguishing the consonant tongue shapes across different tongue length conditions, a linear mixed model (LMM) was performed (Baayen, 2008), separately for each index. Random intercept by speaker, as well as random by Consonant and by Condition slope for individual speakers were included in each LMM. The factor in the model was an interaction variable including all combinations of Consonant (/s/ and /ʃ/ for the alveolar-postalveolar dataset; /k/, /t/, /l/ and /r/ for the alveolar-velar dataset) and Condition (four conditions: Full, Tm, Bm, and T&Bm). The result of the model was subjected to a pairwise comparison procedure, including all combinations of Consonant and Condition. The multiple comparison procedure was implemented in *lsmeans* package (the degrees of freedom were established using the Kenward-Roger approximation), with a Tukey correction (Lenth, 2016). For the alveolar-postalveolar dataset, there were 28 pairwise comparisons for each index in each vowel context: 16 between each consonant in each condition and a different consonant in each condition, and a further 12, between each consonant in each condition and the same consonant in a different condition. For the alveolar-velar dataset, 120 comparisons were carried out for each index: 96 across-consonant combinations, and 24 within-consonant combinations across conditions.

If a given index yielded a significant difference between the two consonants in the Full condition, then it was concluded that the index was suitable for capturing the relevant difference for head-stabilised data, and the index was then included in a further comparison. The results of the other comparisons were used to determine the order of including the indices in a series of generalised linear mixed models (GLMMs), for identifying the optimal combination of indices for characterising the relevant difference in tongue shape. The order depended on how successful each index was in differentiating between the consonants regardless of condition. Specifically, the larger the number of significant across-consonant differences for a given index, the earlier the index was included in the model. GLMMs were run separately for the two datasets (for the alveolar-postalveolar dataset, separately for each vowel context). Consonant was the dependent variable, a random intercept by speaker was included, and the independent variables were Condition (four levels) and one articulatory index in the first model, with each successive model adding one more articulatory index. This process stopped once the comparison of two successive models showed that the addition of another index did not significantly improve the model fit.

[2] For all three new indices, the rightmost point was chosen to be located further back along the tongue curve than the end point, in order to minimise any influence on the measurements from potential issues with imaging the tongue tip.

**Table 1**
Results of multiple comparisons from LMMs on four indices, for the alveolar-postalveolar dataset in the context of /a/. Positive *t* ratio values mean that the mean index value was greater for the postalveolar consonant. The values in bold represent significant differences.

| | $LOC_{a\text{-}i}$ | | TF1 | | TF2 | | TF3 | |
|---|---|---|---|---|---|---|---|---|
| | *t* ratio | *p* | *t* ratio | *p* | *t* ratio | *p* | *t* ratio | *p* |
| /ʃ/$_{Full}$ – /s/$_{Full}$ | **8.524** | **<0.001** | **8.063** | **0.002** | **6.100** | **0.008** | **7.846** | **0.002** |
| /ʃ/$_{Full}$ – /s/$_{Tm}$ | **5.467** | **0.022** | 3.332 | 0.160 | **4.667** | **0.045** | **5.833** | **0.017** |
| /ʃ/$_{Full}$ – /s/$_{Bm}$ | **6.212** | **0.010** | **7.739** | **0.004** | **7.472** | **0.005** | **8.052** | **0.004** |
| /ʃ/$_{Full}$ – /s/$_{T\&Bm}$ | 3.716 | 0.108 | 3.931 | 0.091 | **5.464** | **0.022** | **5.951** | **0.015** |
| /ʃ/$_{Tm}$ – /s/$_{Full}$ | **7.013** | **0.005** | **7.175** | **0.007** | 2.780 | 0.270 | 3.486 | 0.136 |
| /ʃ/$_{Tm}$ – /s/$_{Tm}$ | **8.589** | **<0.001** | **5.841** | **0.011** | 3.416 | 0.122 | 4.124 | 0.052 |
| /ʃ/$_{Tm}$ – /s/$_{Bm}$ | 3.639 | 0.116 | **4.877** | **0.039** | 3.285 | 0.166 | 3.363 | 0.155 |
| /ʃ/$_{Tm}$ – /s/$_{T\&Bm}$ | 2.686 | 0.297 | **6.116** | **0.013** | 3.059 | 0.200 | 3.434 | 0.139 |
| /ʃ/$_{Bm}$ – /s/$_{Full}$ | **5.058** | **0.031** | **5.943** | **0.016** | **5.269** | **0.027** | **6.707** | **0.009** |
| /ʃ/$_{Bm}$ – /s/$_{Tm}$ | 3.486 | 0.137 | 2.549 | 0.340 | 3.816 | 0.100 | 4.345 | 0.062 |
| /ʃ/$_{Bm}$ – /s/$_{Bm}$ | **8.225** | **<0.001** | **6.233** | **0.008** | **7.429** | **0.003** | **9.197** | **0.001** |
| /ʃ/$_{Bm}$ – /s/$_{T\&Bm}$ | **4.591** | **0.047** | 2.992 | 0.222 | 4.244 | 0.067 | 4.396 | 0.059 |
| /ʃ/$_{T\&Bm}$ – /s/$_{Full}$ | 3.164 | 0.189 | **10.660** | **0.001** | 4.404 | 0.058 | **5.582** | **0.021** |
| /ʃ/$_{T\&Bm}$ – /s/$_{Tm}$ | 2.774 | 0.275 | **5.820** | **0.018** | **4.568** | **0.049** | **6.627** | **0.009** |
| /ʃ/$_{T\&Bm}$ – /s/$_{Bm}$ | 3.405 | 0.149 | **8.113** | **0.003** | **4.948** | **0.036** | **5.074** | **0.033** |
| /ʃ/$_{T\&Bm}$ – /s/$_{T\&Bm}$ | **9.651** | **<0.001** | **7.186** | **0.004** | **4.845** | **0.027** | **6.190** | **0.007** |

**Table 2**
Results of multiple comparisons from LMMs on four indices, for the alveolar-postalveolar dataset in the context of /i/. Positive *t* ratio values mean that the mean index value was greater for the postalveolar consonant. The values in bold represent significant differences.

| | $LOC_{a\text{-}i}$ | | TF1 | | TF2 | | TF3 | |
|---|---|---|---|---|---|---|---|---|
| | *t* ratio | *p* | *t* ratio | *p* | *t* ratio | *p* | *t* ratio | *p* |
| /ʃ/$_{Full}$ – /s/$_{Full}$ | **6.043** | **0.009** | **5.892** | **0.012** | **5.533** | **0.009** | **6.248** | **0.007** |
| /ʃ/$_{Full}$ – /s/$_{Tm}$ | **6.958** | **0.007** | 3.484 | 0.138 | **5.117** | **0.029** | 4.070 | 0.079 |
| /ʃ/$_{Full}$ – /s/$_{Bm}$ | 3.823 | 0.100 | **8.440** | **0.002** | **5.147** | **0.028** | **6.356** | **0.011** |
| /ʃ/$_{Full}$ – /s/$_{T\&Bm}$ | **5.885** | **0.017** | 3.010 | 0.217 | **4.965** | **0.031** | 4.130 | 0.074 |
| /ʃ/$_{Tm}$ – /s/$_{Full}$ | 4.156 | 0.074 | **6.227** | **0.012** | 2.397 | 0.386 | 3.652 | 0.114 |
| /ʃ/$_{Tm}$ – /s/$_{Tm}$ | **6.561** | **0.006** | **5.104** | **0.023** | **4.264** | **0.038** | **4.516** | **0.037** |
| /ʃ/$_{Tm}$ – /s/$_{Bm}$ | 2.897 | 0.244 | **5.725** | **0.018** | 2.488 | 0.355 | 3.044 | 0.209 |
| /ʃ/$_{Tm}$ – /s/$_{T\&Bm}$ | 4.348 | 0.063 | 4.406 | 0.056 | 3.169 | 0.173 | 4.013 | 0.077 |
| /ʃ/$_{Bm}$ – /s/$_{Full}$ | **5.595** | **0.019** | 3.880 | 0.095 | **5.361** | **0.023** | **5.646** | **0.020** |
| /ʃ/$_{Bm}$ – /s/$_{Tm}$ | 3.760 | 0.105 | 2.468 | 0.366 | **5.055** | **0.032** | 3.609 | 0.123 |
| /ʃ/$_{Bm}$ – /s/$_{Bm}$ | **5.070** | **0.022** | **5.358** | **0.018** | **6.534** | **0.003** | **6.614** | **0.005** |
| /ʃ/$_{Bm}$ – /s/$_{T\&Bm}$ | **5.865** | **0.017** | 2.092 | 0.511 | **4.793** | **0.038** | 3.569 | 0.127 |
| /ʃ/$_{T\&Bm}$ – /s/$_{Full}$ | **6.755** | **0.008** | **7.008** | **0.007** | 3.752 | 0.101 | **5.206** | **0.028** |
| /ʃ/$_{T\&Bm}$ – /s/$_{Tm}$ | **7.457** | **0.004** | **5.776** | **0.018** | **5.320** | **0.023** | **5.195** | **0.029** |
| /ʃ/$_{T\&Bm}$ – /s/$_{Bm}$ | 3.959 | 0.087 | **6.578** | **0.010** | 3.666 | 0.112 | 4.413 | 0.057 |
| /ʃ/$_{T\&Bm}$ – /s/$_{T\&Bm}$ | **6.647** | **0.006** | **5.225** | **0.021** | **4.599** | **0.025** | **5.236** | **0.018** |

## 3. Results

### 3.1. Alveolar-postalveolar dataset

The LMM results showed that in the Full condition, DEI was the only index that did not yield a significant difference between /ʃ/ and /s/ in either vowel context (/ʃ/$_{Full}$[3] versus /s/$_{Full}$: *t* ratio = 2.95, $p = 0.23$ in the context of /a/; *t* ratio = 2.38, $p = 0.39$ in the context of /i/). Further comparisons then did not include DEI, focussing on the other four indices. The results of across-consonant pairwise comparisons for the two sibilants are reported in Table 1 for the context of /a/, and in Table 2 for the context of /i/. There were no significant differences in within-consonant pairwise comparisons, except one significant difference for TF3 in the context of /a/ (/ʃ/$_{T\&Bm}$ larger than /ʃ/$_{Right}$; *t* ratio = 4.27, $p = 0.021$).

For the /ʃ/-/s/ pair in the context of /a/, TF1 differentiated between the consonants in 12 out of 16 possible combinations of consonant and condition. This was followed by TF3 (ten combinations), then by $LOC_{a\text{-}i}$ and TF2 (nine each). $LOC_{a\text{-}i}$ was ranked higher than TF2 because $LOC_{a\text{-}i}$ was the only one of the four indices to have a significant result for the /ʃ/$_{Bm}$ – /s/$_{T\&Bm}$ pair. For the /ʃ/-/s/ pair in the context of /i/, TF2 had the maximum number of combinations where the two consonants were significantly different from each other (11), followed by $LOC_{a\text{-}i}$ and TF1 (ten each), and TF3 (eight). In the context of /i/, TF1 was ranked higher than $LOC_{a\text{-}i}$ because the former index was the only one out of four indices to differentiate between the two consonants in three different pairs of conditions (/ʃ/$_{Tm}$ – /s/$_{Full}$; /ʃ/$_{Tm}$ – /s/$_{Bm}$; /ʃ/$_{T\&Bm}$ – /s/$_{Bm}$). For each vowel context, the indices were included in the GLMMs in the order described above. Table 3 shows the results of the likelihood ratio tests comparing the successive models for each vowel context. In each of the two vowel contexts, adding together up to three indices led to a significant improvement in the model fit, and adding a fourth index did not lead to any further improvement. The three indices that together constituted an optimal combination for differentiating between the two fricatives are represented graphically in Fig. 7 (using the package *scatterplot3d*: Ligges and Mächler, 2003), separately for the two vowel contexts. For each vowel context, the values for the index that was the most successful in differentiating the consonants are plotted along the horizontal axis, the values for the next successful index are plotted along the axis that is shown as a diagonal line in Fig. 7. Finally, the values for third successful index are plotted along the vertical axis. Visual inspection of the figure

[3] In the reporting of the results, the subscript indication of the condition is used, as follows: "/s/$_{Full}$" refers to the set of /s/ curves in the Full condition, "/k/$_{Tm}$" refers to the set of /k/ curves in the Tm condition, and so on.

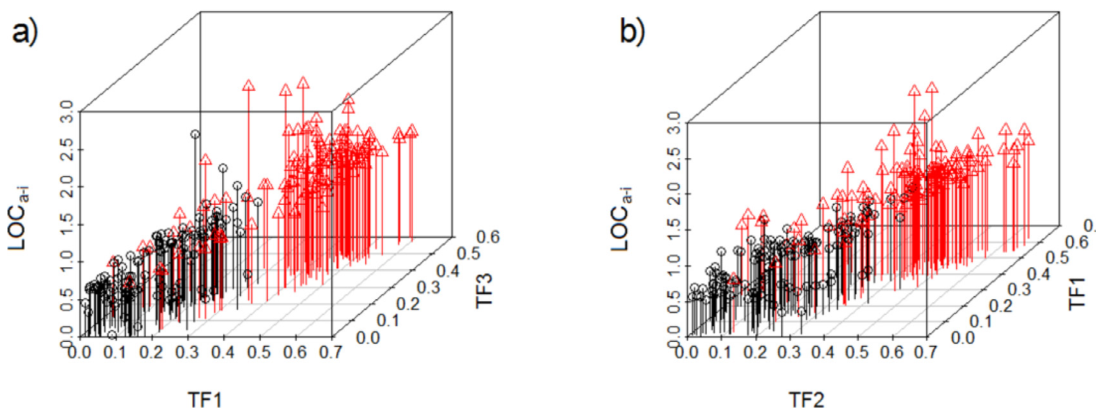


**Fig. 7.** Values of the three indices which, when included in the GLMM for the alveolar-postalveolar contrast, led to a significant improvement in the model fit: a) the context of /a/, b) the context of /i/; circles represent /s/, and triangles represent /ʃ/. The values are pooled across conditions.

**Table 3**
Results of the likelihood ratio tests comparing successive GLMMs for each vowel context, for the /ʃ/-/s/ contrast. The second to the fifth columns of the table are labelled as per the output of the likelihood ratio tests, as follows: "AIC" stands for "Akaike Information Criterion"; "Chisq Chi" is the Chi-square value; in the "Df" column, each value represents the change in degrees of freedom between the given model and the model in the row immediately above; the last column includes *p* values for the model comparisons. Values in bold represent the models where adding an index as a predictor resulted in a significant improvement of the model fit.

| /ʃ/ versus /s/, context of /a/ | AIC | Chisq Chi | Df | Pr(>Chisq) |
|---|---|---|---|---|
| TF1 | 108.0 | | | |
| TF1 + TF3 | 100.6 | 9.40 | 1 | **0.002** |
| TF1 + TF3 + $LOC_{a\text{-}i}$ | 90.1 | 12.50 | 1 | **<0.001** |
| TF1 + TF3 + $LOC_{a\text{-}i}$ + TF2 | 90.5 | 1.62 | 1 | 0.204 |
| /ʃ/ versus /s/, context of /i/ | | | | |
| TF2 | 172.8 | | | |
| TF2 + TF1 | 111.2 | 63.57 | 1 | **<0.001** |
| TF2 + TF1 + $LOC_{a\text{-}i}$ | 39.5 | 73.70 | 1 | **<0.001** |
| TF2 + TF1 + $LOC_{a\text{-}i}$ + TF3 | 40.5 | 1.04 | 1 | 0.309 |

shows that the two consonants are noticeably differentiated along the horizontal dimension, with most of the values for /s/ clustering towards the left of this axis, in both vowel contexts. Examining the distribution along the diagonal axis, again, shows that the values for each of the consonants are clustered towards the opposing sides of the axis, with the alveolar fricative values in the lower end. In the vertical dimension, the differences are somewhat less prominent visually than in the other two dimensions, but the values for the two consonants are still distributed in the same direction, with /s/ having generally smaller values than /ʃ/.

### *3.2. Alveolar-velar dataset*

In the Full condition comparisons in the context of /a/, two indices did not yield significantly greater values for /k/ than for each alveolar consonant. Specifically, for $LOC_{a\text{-}i}$ and TF1, the difference between /k/ and /t/ did not reach significance. Tables 4 and 5 show the results of the pairwise across-consonant comparisons for DEI and TF2, respectively. These two indices complemented each other, together differentiating between all alveolar-velar combinations, across different conditions. The results for TF3 were very similar to those for TF2, with the exception of three differences that were significant for TF2 but not for TF3: $/k/_{Full} - /t/_{Tm}$, $/k/_{Full} - /t/_{T\&Bm}$, and $/k/_{Tm} - /l/_{Full}$. In within-consonant pairwise comparisons, there were no significant differences for DEI and TF2, except the following pairs: DEI $/k/_{Full}$ larger than $/k/_{Bm}$, $t$ ratio = 5.33, $p = 0.027$; DEI $/t/_{Full}$ larger than $/t/_{T\&Bm}$,

**Table 4**
Results of multiple comparisons between /k/ and alveolar consonants on DEI, for the context of /a/. "ALV" stands for "alveolar", with results for individual alveolar consonants included in separate columns of the table. Positive *t* ratio values mean that the mean index value was greater for the velar consonant. The values in bold represent significant differences.

| | /t/ | | /l/ | | /r/ | |
|---|---|---|---|---|---|---|
| | *t* ratio | *p* | *t* ratio | *p* | *t* ratio | *p* |
| $/k/_{Full} - ALV_{Full}$ | **10.642** | **0.002** | **9.115** | **0.004** | **15.964** | **<0.001** |
| $/k/_{Full} - ALV_{Tm}$ | **9.567** | **0.004** | **8.530** | **0.007** | **12.895** | **0.001** |
| $/k/_{Full} - ALV_{Bm}$ | **10.747** | **0.002** | **10.112** | **0.003** | **12.810** | **0.001** |
| $/k/_{Full} - ALV_{T\&Bm}$ | **11.305** | **0.002** | **11.034** | **0.002** | **15.738** | **<0.001** |
| $/k/_{Tm} - ALV_{Full}$ | **10.119** | **0.003** | **10.147** | **0.003** | **13.547** | **0.001** |
| $/k/_{Tm} - ALV_{Tm}$ | **12.355** | **0.001** | **11.528** | **0.001** | **18.648** | **<0.001** |
| $/k/_{Tm} - ALV_{Bm}$ | **8.734** | **0.006** | **9.027** | **0.005** | **9.474** | **0.004** |
| $/k/_{Tm} - ALV_{T\&Bm}$ | **11.302** | **0.002** | **12.951** | **0.001** | **14.669** | **<0.001** |
| $/k/_{Bm} - ALV_{Full}$ | **6.687** | **0.020** | 5.136 | 0.064 | **10.251** | **0.002** |
| $/k/_{Bm} - ALV_{Tm}$ | **5.843** | **0.038** | 5.209 | 0.062 | **7.754** | **0.010** |
| $/k/_{Bm} - ALV_{Bm}$ | **8.381** | **0.006** | **5.974** | **0.031** | **11.685** | **0.001** |
| $/k/_{Bm} - ALV_{T\&Bm}$ | **7.734** | **0.010** | **6.544** | **0.023** | **10.720** | **0.002** |
| $/k/_{T\&Bm} - ALV_{Full}$ | **9.944** | **0.003** | **7.523** | **0.012** | **13.850** | **<0.001** |
| $/k/_{T\&Bm} - ALV_{Tm}$ | **9.264** | **0.004** | **7.597** | **0.012** | **12.200** | **0.001** |
| $/k/_{T\&Bm} - ALV_{Bm}$ | **10.149** | **0.003** | **8.104** | **0.008** | **11.351** | **0.001** |
| $/k/_{T\&Bm} - ALV_{T\&Bm}$ | **11.426** | **0.001** | **9.760** | **0.003** | **15.725** | **<0.001** |

**Table 5**
Results of multiple comparisons between /k/ and alveolar consonants on TF2, for the context of /a/. "ALV" stands for "alveolar", with results for individual alveolar consonants included in separate columns of the table. Positive *t* ratio values mean that the mean index value was greater for the velar consonant. The values in bold represent significant differences.

| | /t/ | | /l/ | | /r/ | |
|---|---|---|---|---|---|---|
| | *t* ratio | *p* | *t* ratio | *p* | *t* ratio | *p* |
| $/k/_{Full} - ALV_{Full}$ | **6.581** | **0.007** | **8.034** | **0.001** | **5.689** | **0.018** |
| $/k/_{Full} - ALV_{Tm}$ | **6.618** | **0.010** | **9.573** | **0.001** | **7.557** | **0.003** |
| $/k/_{Full} - ALV_{Bm}$ | 5.010 | 0.059 | 5.195 | 0.052 | 2.767 | 0.448 |
| $/k/_{Full} - ALV_{T\&Bm}$ | **7.151** | **0.007** | **7.482** | **0.005** | 4.148 | 0.115 |
| $/k/_{Tm} - ALV_{Full}$ | 4.305 | 0.100 | **5.092** | **0.048** | 3.287 | 0.275 |
| $/k/_{Tm} - ALV_{Tm}$ | 4.536 | 0.059 | **6.225** | **0.008** | 4.584 | 0.058 |
| $/k/_{Tm} - ALV_{Bm}$ | 3.271 | 0.289 | 3.434 | 0.252 | 1.425 | 0.957 |
| $/k/_{Tm} - ALV_{T\&Bm}$ | 4.722 | 0.067 | 4.662 | 0.073 | 2.123 | 0.719 |
| $/k/_{Bm} - ALV_{Full}$ | **8.926** | **0.002** | **13.100** | **<0.001** | **10.864** | **<0.001** |
| $/k/_{Bm} - ALV_{Tm}$ | **7.740** | **0.006** | **12.175** | **<0.001** | **11.374** | **<0.001** |
| $/k/_{Bm} - ALV_{Bm}$ | **9.535** | **0.001** | **10.824** | **<0.001** | **6.723** | **0.007** |
| $/k/_{Bm} - ALV_{T\&Bm}$ | **9.832** | **0.001** | **12.858** | **<0.001** | **8.580** | **0.001** |
| $/k/_{T\&Bm} - ALV_{Full}$ | **8.040** | **0.004** | **9.818** | **0.001** | **7.814** | **0.004** |
| $/k/_{T\&Bm} - ALV_{Tm}$ | **8.036** | **0.004** | **11.335** | **<0.001** | **10.033** | **0.001** |
| $/k/_{T\&Bm} - ALV_{Bm}$ | **6.332** | **0.021** | **6.571** | **0.018** | 4.143 | 0.131 |
| $/k/_{T\&Bm} - ALV_{T\&Bm}$ | **9.064** | **0.001** | **9.837** | **<0.001** | **6.432** | **0.009** |

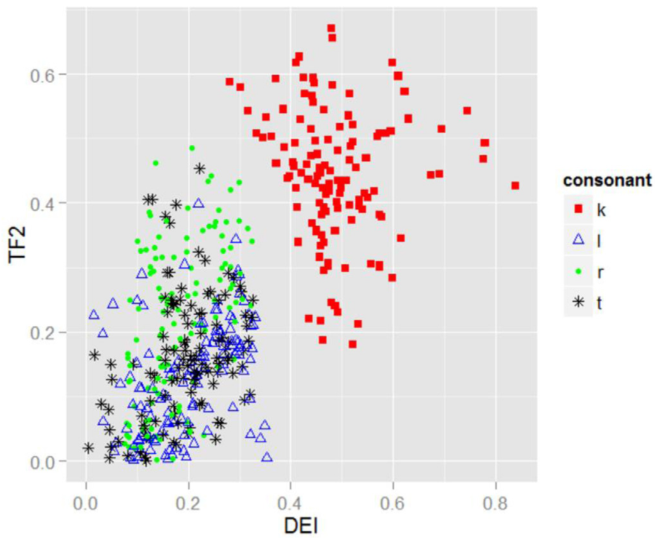


**Fig. 8.** Values of the two indices that accounted for all of the consonant-related variation in the alveolar-velar dataset. The values are pooled across conditions.

*t* ratio = 5.01, *p* = 0.014; DEI /l/$_{\text{Full}}$ larger than /l/$_{\text{T\&Bm}}$, *t* ratio = 4.87, *p* = 0.017; TF2 /k/$_{\text{T\&Bm}}$ larger than /k/$_{\text{Tm}}$, *t* ratio = 5.93, *p* = 0.001.

A GLMM with DEI as a predictor was run on the alveolar-velar dataset, and another model was then run on the same data, including both DEI and TF2. A comparison of the two models was not possible because the second model did not converge, due to the fact that the combination of these two indices differentiated between all the data points, without overlap. Fig. 8 shows the values of the two indices that together accounted for all of the consonant-related variation.

## 4. Discussion

The study identified several measures of tongue shape that differentiated between contrasting places of articulation for two lingual contrasts in English: alveolar-postalveolar, and alveolar-velar. For each of the contrasts, combinations of measures were found to better distinguish between place of articulation contrasts than individual measures. Robust place of articulation distinctions were observed regardless of differences in the length of the imaged tongue curve, when combinations of measures were used. Since all five measures discussed in the paper do not require that the tongue curves for multiple repetitions are located in the same coordinate space, combinations of these measures allow for a comparison of tongue curves across repetitions in recordings without head-to-transducer stabilisation (such as recordings of young children and people with mobility issues), as well as generally across speakers, and across multiple recording sessions for the same speaker. While this study focussed on establishing *if* tongue shapes for contrasting sounds are different, rather than *how* different they are, this latter point could be addressed in future work with all five indices, to quantify the extent of differences between tongue shapes for contrasting targets, and compare it across groups of speakers. The present study only analysed productions by adults with typical speech, and further work is required to test the performance of the indices on other types of data, such as productions by typically developing children. This is the focus of ongoing and further work, which requires collecting previously unavailable datasets of typical child speech and analysing tongue shape contrasts in different age groups. Generally, in future instrumental studies of the development of phonological oppositions in young children's speech, it would be useful to include ultrasound data on tongue shape, in addition to acoustic data that have often been analysed in such studies to date (e.g., Macken and Barton, 1980; Scobbie et al., 2000; Edwards and Beckman, 2008; Munson et al., 2010; Holliday et al., 2015).

For the alveolar-postalveolar distinction, $\text{LOC}_{\text{a-i}}$ and three new measures (TF1, TF2 and TF3) were shown to distinguish between contrasting places of articulation for the sibilant consonants in the Full condition, i.e., for the original, full-length tongue curves. The finding for $\text{LOC}_{\text{a-i}}$ was expected, since it has previously been used to differentiate /s/ and /ʃ/ tongue shapes (in particular, for the larger corpus incorporating the dataset from the present study; Zharkova, 2016). The result from the three radius-based measures indicates that each one of them can also be used to capture the alveolar-postalveolar distinction in head-to-transducer stabilised data on tongue shape. When all combinations of consonant and condition were taken into account, each of the four indices did not distinguish between the two sibilants in every pair of conditions, and different indices complemented each other in distinguishing between the consonants across pairs of conditions. Different combinations of indices worked better in different vowel contexts, suggesting that in further studies of tongue shape differences across alveolar and postalveolar places of articulation, vowel context needs to be taken into account when analysing the data.

For analysing the alveolar-velar distinction, the present paper included three different alveolar consonants, so the place of articulation difference was assessed beyond using the same manner of articulation, such as would be the case if only /k/-/t/ opposition was analysed. DEI was, as expected, found to distinguish between the tongue shapes for each of the alveolar consonants and the tongue shape for the velar stop, in the Full condition. Two other indices, TF2 and TF3, also differentiated between each of the alveolars and the velar, in that condition. DEI and TF2 together accounted for all alveolar-velar differences in tongue shape, and the two indices complemented each other in differentiating between the places of articulation across conditions. The fact that the three alveolar consonants patterned together, all contrasting with /k/, is relevant for potential clinical applications, for example, in relation to the phonological process of realising liquids as [w]. The tongue shape of the labial-velar approximant is very similar to that of velar stops (see quantifications based on ultrasound data, reported by Dawson et al., 2016). Therefore, in addition to analysing alveolar-velar stop contrasts, being able to distinguish between target tongue shapes of alveolars and velars makes it possible to assess whether the phonological process of gliding as judged perceptually may involve underlying lingual targets that do not necessarily lend themselves to auditory identification.

While the two sibilants were differentiated in both vowel contexts (/a/ and /i/), the high vowel context was not suitable for analysing tongue shape differences for the velar stop versus the alveolar consonants. Indeed, the high front vowel context was not included in the alveolar-velar dataset because in this context none of the five indices distinguished consistently across contrasting places of articulation (cf. Recasens and Rodríguez, 2018b, showing stronger effects on consonantal lingual articulation from the vowel context of /i/ than from the vowel context of /a/ in Catalan, using midsagittal ultrasound data). This has further implications related to choosing the phonological environment for targeting certain contrasts, both for future research studies and for clinical applications, i.e., for selecting linguistic materials for ultrasound biofeedback therapy. In the high front vowel context, other measures would be required to capture the relevant articulatory differences, such as an acoustic measure of centroid frequency for [k] versus [t] (see, e.g., Johnson et al., 2018). It would be useful to include more vowel contexts in further studies, to establish whether the /k/-/t/ difference in tongue shape would be captured by the indices in other contexts. The context of /u/ in Scottish English would likely provide little additional information to that reported here for the context of /i/, because those two vowels are quite similar to each other in tongue shape in this variety of English (see, e.g., Zharkova et al., 2012). Consonant oppositions in other vowel contexts, such as open-mid and close-mid vowels (both front and back) would be interesting to analyse, and we could hypothesise that the velar-alveolar distinction might hold in at least open-mid vowel contexts.

Since both lingual contrasts of interest in this paper primarily involve shape differences in the front two thirds of the tongue contour (see Figs. 1–3), the measurements used in the paper were focussed on differentiating tongue shapes in that region of the curve. DEI, as expected, reflected differences in the extent of overall tongue bunching between velar and alveolar tongue shapes in the low vowel context, while differences between the postalveolar and the alveolar fricative in tongue predorsum bunching (more for the former consonant) in both low and high vowel contexts were captured by $LOC_{a\text{-}i}$. More fine grained differences in tongue shape were reflected by the three new indices, with each index encompassing a specific region of the curve. Below it is reviewed how local shape differences across contrasting consonants were represented by each of the three new indices, and suggestions are made regarding the applicability of these indices to measuring tongue contrasts beyond those discussed here.

In the present paper, TF1 was useful in distinguishing between /s/ and /ʃ/. The calculations for this index involved the region of the tongue curve comparable to what Recasens and Rodríguez (2016) referred to as the palatal zone, i.e., further back along the tongue curve than the region serving as primary articulator for alveolar consonants, but generally further forward than that involved in making constrictions for the production of velars. Thus, TF1 reflected the differences in tongue shape that were due to predorsum raising for the postalveolar fricative /ʃ/. This index was much less informative on the velar-alveolar distinction because the relevant stretch of the curve did not particularly differ in shape for /k/ and the alveolars, with the main difference in shape occurring further back along the tongue curve, beyond the reach of TF1. By contrast, TF2 and TF3 were both centred more posteriorly along the tongue curve than TF1. This enabled those indices to capture changes in curvature occurring further back along the midsagittal tongue contour, such as those for /k/, compared with the alveolars. It is interesting that TF2 and TF3 performed differently for the /s/-/ʃ/ opposition in the different vowel contexts, with the former carrying more weight in the context of the high front vowel, and the latter being more relevant in the context of the low vowel. With relatively little influence from the low vowel on midsagittal tongue shape at mid-fricative, the change in shape across the two fricatives encompasses a relatively wide area in the front half of the tongue curve. This change is accordingly better captured by TF3 than by TF2, since TF3 spans over a wider region of the curve. The narrower region captured by TF2, on the opposite, becomes more relevant when the difference between the two fricatives becomes partly obscured by the global change in tongue position in anticipation of the vowel /i/, particularly prominent for /s/ (see, e.g., Zharkova et al., 2018). While in the present study the new indices were applied to specific consonant contrasts and vocalic contexts, the indices would also be likely to reflect other linguistically relevant shape distinctions in the front two thirds of the tongue curve (note that the indices were not designed to capture shape differences at the back of the curve). It could be hypothesised that in further studies with other datasets, useful information may be obtained on midsagittal tongue shapes for other front lingual places of consonant articulation, such as palatals and alveolo-palatals, for front vowels of different degrees of jaw opening, and for contrasting consonants in other vowel contexts, such as close-mid and open-mid front vowels.

An important advantage of the new measures introduced in this paper is that they focus on very specific parts of the tongue curve, with the value of each index related to linguistically relevant articulations, and that at the same time the indices can be applied to non-head-stabilised ultrasound data. Other existing measures applicable to non-stabilised data are either not as localised within the curve (e.g., DEI and $LOC_{a\text{-}i}$; other ratio-based tongue shape measures described in Aubin and Ménard, 2006, and Zharkova, 2013) or not focussed on any specific location within the curve, with measurements not easily related to linguistic distinctions (e.g., a discrete Fourier transform of a tongue shape described in Dawson et al., 2016; see also Abakarova et al., 2018, reporting results from using such measurements). Existing measures that are able to compare specific parts of the tongue curve do so at the cost of requiring head-stabilised data (e.g., single point measurements: Ménard et al., 2012; Noiray et al., 2018; Abakarova et al., 2018; Rubertus and Noiray, 2018; regions of the tongue curve: Recasens and Rodríguez, 2016; 2018a; 2018b). Since non-stabilised ultrasound recordings induce variability in the data, there is a challenge in controlling for that variability while being able to identify the relevant linguistic distinctions between targets of interest. In the present study, this was achieved by finding optimal combinations of measures to differentiate between contrasting consonant targets, and combining measures was reported to be more powerful than using individual measures, when all tongue length conditions were taken into account.

Various combinations of the five measures discussed in this work can be used to address other lingual contrasts, including those involving different manners of articulation to those reported here (cf. Liker and Gibbon, 2012, an electropalatographic and perceptual study of Croatian postalveolar and palatal affricates). The contrasts analysed in this study were from English, and it is likely that different patterns would have been observed if data from other languages had been used, particularly in the case of sibilant fricative distinctions in place of articulation (see, e.g., Gordon et al., 2002). Based on the findings from this study, we could expect that different combinations of indices would provide optimal for differentiating between contrasting consonants, depending on the language or accent studied.

## Declaration of competing interest

None.


## Acknowledgments

The author wishes to thank Bill Hardcastle for his comments on this work. Thanks are due to Véronique Delvaux and two anonymous reviewers for helpful suggestions. This work was supported by two grants from the Economic and Social Research Council to the author: RES-000-22-4075 and ES/K002597/1.


## References


Abakarova, D., Iskarous, K., Noiray, A., 2018. Quantifying lingual coarticulation in German using mutual information: an ultrasound study. J. Acoust. Soc. Am. 144, 897–907.

Articulate Instruments Ltd., 2012. Articulate Assistant Advanced Ultrasound Module User Guide: Version 2.14. Articulate Instruments Ltd, Edinburgh, UK.

Aubin, J., Ménard, L., 2006. Compensation for a labial perturbation: an acoustic and articulatory study of child and adult French speakers. In: Yehia, H.C., Demolin, D., Laboissière, R. (Eds.), Proceedings of the Seventh International Seminar on Speech Production, Ubatuba, Brazil, pp. 209–216.

Baayen, R.H., 2008. Analyzing Linguistic Data. A Practical Introduction to Statistics Using R. Cambridge University Press, Cambridge.

Beckman, M.E., Yoneyama, K., Edwards, J., 2003. Language-specific and language-universal aspects of lingual obstruent productions in Japanese-acquiring children. J. Phon. Soc. Jpn. 7, 18–28.

Bressmann, T., Thind, P., Bollig, C.M., Uy, C., Gilbert, R.W., Irish, J.C., 2005. Quantitative three-dimensional ultrasound analysis of tongue protrusion, grooving and symmetry: data from twelve normal speakers and a partial glossectomee. Clin. Linguist. Phon. 19, 573–588.

Davidson, L., 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. J. Acoust. Soc. Am. 120, 407–415.

Dawson, K.M., Tiede, M.K., Whalen, D.H., 2016. Methods for quantifying tongue shape and complexity using ultrasound imaging. Clin. Linguist. Phon. 30, 328–344.

Dodd, B., 2013. Differential Diagnosis and Treatment of Children with Speech Disorder, second ed. John Wiley & Sons, London, UK.

Edwards, J., Beckman, M.E., 2008. Methodological questions in studying consonant acquisition. Clin. Linguist. Phon. 22, 937–956.

Edwards, J., Gibbon, F., Fourakis, M., 1997. On discrete changes in the acquisition of the alveolar/velar stop consonant contrast. Lang. Speech 40, 203–210.

Gibbon, F., Lee, A., 2017a. Preface to the special issue on covert contrasts. Clin. Linguist. Phon. 31, 1–3.

Gibbon, F.E., Lee, A., 2017b. Electropalatographic (EPG) evidence of covert contrasts in disordered speech. Clin. Linguist. Phon. 31, 4–20.

Gick, B., Bacsfalvi, P., Bernhardt, B.M., Oh, S., Stolar, S., Wilson, I., 2008. A motor differentiation model for liquid substitutions: English /r/ variants in normal and disordered acquisition. Proc. Meet. Acoust. 1, 1–9.

Gordon, M., Barthmaier, P., Sands, K., 2002. A cross-linguistic acoustic study of voiceless fricatives. J. Int. Phon. Assoc. 32, 141–174.

Hewlett, N., 1988. Acoustic properties of /k/ and /t/ in normal and phonologically disordered speech. Clin. Linguist. Phon. 2, 29–45.
Holliday, J.J., Reidy, P.F., Beckman, M.E., Edwards, J., 2015. Quantifying the robustness of the English sibilant fricative contrast in children. J. Speech Lang. Hear. Res. 58, 622–637.
Johnson, A.A., Reidy, P.F., Edwards, J.R., 2018. Quantifying robustness of the /t/-/k/ contrast using a single, static spectral feature. J. Acoust. Soc. Am. 144, EL105–EL111.
Kansy, K., Hoffmann, J., Mistele, N., Shavlokhova, V., Bendszus, M., Heiland, S., Krisam, J., Geschwinder, A., Gradl, J., 2018. Visualization and quantification of tongue movement during articulation: is ultrasound a valid alternative to magnetic resonance imaging? J. Craniomaxillofac. Surg. 46, 1924–1933.
Lenth, R.V., 2016. Least-Squares means: the R package lsmeans. J. Stat. Softw. 69, 1–33.
Li, F., Edwards, J., Beckman, M.E., 2009. Contrast and covert contrast: the phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. J. Phon. 37, 111–124.
Ligges, U., Mächler, M., 2003. Scatterplot3d – an R package for visualizing multivariate data. J. Stat. Softw. 8, 1–20.
Liker, M., Gibbon, F., 2012. An EPG and perceptual study of the postalveolar and palatal affricate contrast in standard Croatian. Ital. J. Linguist. 24, 43–64.
Macken, M.A., Barton, D., 1980. The acquisition of the voicing contrast in English: a study of voice onset time in word-initial stop consonants. J. Child Lang. 7, 41–74.
McAllister Byun, T., Buchwald, A., Mizoguchi, A., 2016. Covert contrast in velar fronting: an acoustic and ultrasound study. Clin. Linguist. Phon. 30, 249–276.
Ménard, L., Aubin, J., Thibeault, M., Richard, G., 2012. Comparing tongue shapes and positions with ultrasound imaging: a validation experiment using an articulatory model. Folia Phoniatr. Logop. 64, 64–72.
Munson, B., Edwards, J., Schellinger, S.K., Beckman, M.E., Meyer, M.K., 2010. Deconstructing phonetic transcription: covert contrast, perceptual bias, and an extraterrestrial view of vox humana. Clin. Linguist. Phon. 24, 245–260.
Noiray, A., Abakarova, D., Rubertus, E., Krüger, S., Tiede, M., 2018. How do children organize their speech in the first years of life? Insight from ultrasound imaging. J. Speech Lang. Hear. Res. 61, 1355–1368.
Preston, J., McCabe, P., Tiede, M., Whalen, D.H., 2019. Tongue shapes for rhotics in school-age children with and without residual speech errors. Clin. Linguist. Phon. 33, 334–348.
R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna [computer program] http://www.R-project.org. (Last viewed May 16, 2013).
Recasens, D., Rodríguez, C., 2016. A study of coarticulatory resistance and aggressiveness for front lingual consonants and vowels using ultrasound. J. Phon. 59, 58–75.
Recasens, D., Rodríguez, C., 2018a. Contextual and syllabic effects in heterosyllabic consonant sequences. An ultrasound study. Speech Commun. 96, 150–167.
Recasens, D., Rodríguez, C., 2018b. An ultrasound study of contextual and syllabic effects in consonant sequences produced under heavy articulatory constraint conditions. Speech Commun. 105, 34–52.
Rubertus, E., Noiray, A., 2018. On the development of gestural organization: a cross-sectional study of vowel-to-vowel anticipatory coarticulation. PLoS One 13, e0203562.
Scobbie, J.M., Gibbon, F., Hardcastle, W.J., Fletcher, P., 2000. Covert contrast as a stage in the acquisition of phonetics and phonology. In: Broe, M.B., Pierrehumbert, J.B. (Eds.), Papers in Laboratory Phonology V: Acquisition and the Lexicon. Cambridge University Press, Cambridge, pp. 194–207.
Smit, A.B., 1993. Phonologic error distributions in the Iowa-Nebraska Articulation Norms Project: consonant singletons. J. Speech Hear. Res. 36, 533–547.
Stone, M., 2010. Laboratory techniques for investigating speech articulation. In: Hardcastle, W.J., Laver, J., Gibbon, F.E. (Eds.), The Handbook of Phonetic Sciences. Wiley-Blackwell, Chichester, United Kingdom, pp. 9–38.
Stone, M., Morrish, K.A., Sonies, B.C., Shawker, T.H., 1987. Tongue curvature: a model of shape during vowel production. Folia Phoniatr. 39, 302–315.
Tyler, A.A., Figurski, G.R., Langsdale, T., 1993. Relationships between acoustically determined knowledge of stop place and voicing contrasts and phonological treatment progress. J. Speech Hear. Res. 36, 746–759.
Zharkova, N. (2007). Quantification of coarticulatory effects in several Scottish English phonemes using ultrasound. Working Papers, QMU Speech Science Research Centre WP-13, 1–19.
Zharkova, N., 2013. A normative-speaker validation study of two indices developed to quantify tongue dorsum activity from midsagittal tongue shapes. Clin. Linguist. Phon. 27, 484–496.
Zharkova, N., 2016. Ultrasound and acoustic analysis of sibilant fricatives in preadolescents and adults. J. Acoust. Soc. Am. 139, 2342–2351.
Zharkova, N., Gibbon, F.E., Hardcastle, W.J., 2015. Quantifying lingual coarticulation using ultrasound imaging data collected with and without head stabilisation. Clin. Linguist. Phon. 29, 249–265.
Zharkova, N., Gibbon, F.E., Lee, A., 2017. Using ultrasound tongue imaging to identify covert contrasts in children's speech. Clin. Linguist. Phon. 31, 21–34.
Zharkova, N., Hardcastle, W.J., Gibbon, F.E., 2018. The dynamics of voiceless sibilant fricative production in children between seven and thirteen years old: an ultrasound and acoustic study. J. Acoust. Soc. Am. 144, 1454–1466.
Zharkova, N., Hewlett, N., 2009. Measuring lingual coarticulation from midsagittal tongue contours: description and example calculations using English /t/ and /ɑ/. J. Phon. 37, 248–256.
Zharkova, N., Hewlett, N., Hardcastle, W.J., 2012. An ultrasound study of lingual coarticulation in /sV/ syllables produced by adults and typically developing children. J. Int. Phon. Assoc. 42, 193–208.