Searching for Closure: Seeing a Dip

Cornelia J Heyde[1], James M Scobbie[1], Ian Finlayson[1,2]

[1] *Clinical Audiology, Speech and Language (CASL) Research Centre, Queen Margaret University, Edinburgh, UK*
[2] *School of Philosophy, Psychology and Language Sciences (PPLS), Edinburgh University, Edinburgh, UK*

Quantifying lingual kinematics in relation to passive articulators is as crucial and elementary as it is challenging for ultrasound tongue imaging (UTI) research. In UTI, generally only the active tongue is observable, with passive articulatory structures such as the hard and soft palate being invisible almost all of the time. The fact that the tongue can take on various lengths and an almost indefinite set of shapes further accounts for the difficulty in establishing a referent that would allow for inter-speaker comparison. Finding a referent that respects articulatory heterogeneity is a persistent challenge. In the case of a velar stop, for example, how is the constriction found in the image? Frisch (2010) has argued for the value of automatic detection of the location of a constriction based on the shape of the tongue surface as it is deformed by contact, thereby relying on the tongue shape itself. Another approach that avoids external referents is that of Iskarous (2005) who has investigated pivot points to explore patterns in tongue contour deformation in dynamic data.

In the current study we propose a method that uses both dynamic data and movement patterns to establish the location of the constriction. The method serves to identify a referent/measurement vector along which tongue motion during the approach to and movement away from a constriction can be measured speaker-independently. We report the use of this novel technique as applied to velar closures. The resulting measures obtained along the vector can be used to quantify the degree and timing of lingual movement before and after closure, while also identifying the location of the constriction.

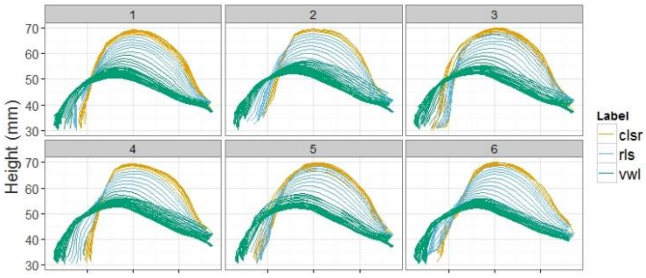

Figure 1 - splined tongue contours (tongue root on the left; tongue tip on the right) for six productions of the same /kɑ/ prompt produced by the same speaker B

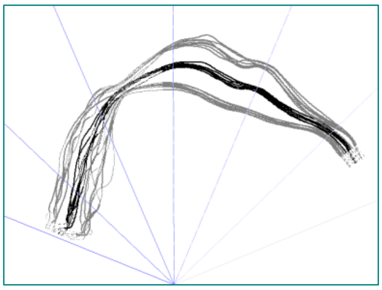

Figure 2 - overlaid mean splines (black) and SDs (grey) for the six productions of /kɑ/ (Figure 1)

The technique takes as its input multiple tokens of /kV/ targets which have been semi-automatically splined for about 700 ms (Figure 1; Articulate Instruments Ltd 2012). A fan-shaped grid of 42 equidistant radial fanlines is superimposed (Figure 2**)**. The polar co-ordinates at which each fanline intersects with the spline are recorded. This allows us to calculate the distance to the surface from a virtual origin located within the ultrasound probe. Distances from the probe to the tongue surface at adjacent fanlines are clearly going to be highly correlated. We plotted these correlations (Pearson's r; Figure 3) for splines that were extracted from the acoustic midpoint of the closure and found they can be used to guide the placement of a measurement vector, a fanline. As expected, there was always an extremely high correlation of the polar distances to the tongue surface of adjacent fanlines as calculated across repetitions of the same phoneme.

We noticed however a 'dip' (in a few cases we observe multiple dips, such as for speakers I and K in the bottom left and bottom right panels of Figure 3) that occurs in the midst of the overall high correlations of each speaker's correlated data. Plotting r for all adjacent fanline pairs along the tongue surface therefore results in high correlations at adjacent splines on the tongue surface, generally speaking, with a dip in correlation of two fanlines. The correlation dips stand for a reduced reliability of the location of the tongue spline for the respective area. In all but one of the cases (cf. speaker A in the upper leftmost panel) the most prominent correlation dips occur relatively central (near fanline 21) to the correlated fanlines of the ultrasound image, which is also where we would expect the tongue to form the palatal constriction in the case of /k/.

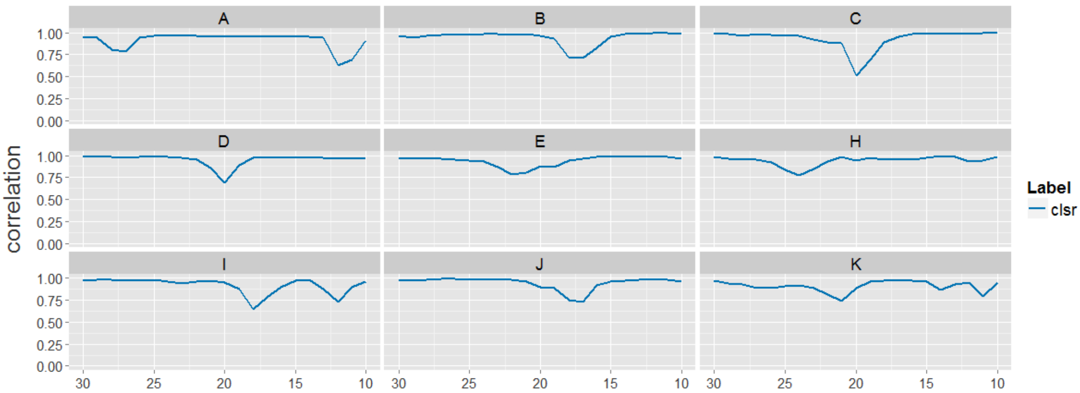

Figure 3 - Pearson's r correlations of adjacent radial fanlines along the tongue surface (from left = posterior to right = anterior) across multiple repetitions of /kV/ produced by 9 speakers (A – K)

In a previous study also on the formation of velar closure (also including the data for the current study) we have semi-manually established the fanline along which the extent of lingual movement is greatest. Interestingly, we found a meaningful overlap of those semi-manually established fanlines and the fanlines marked by the correlation dip in the current study. The systematic occurrence of the dips in addition to the clear overlap of their location and that of the semi-manually found fanlines is intriguing.

Together this indicates that dips are more than random occurrences. Dips are likely to be related to the closing gesture from which the splines were extracted. They may be particularly useful in the study of motor control as they may indicate: (1) the location of the tongue at closure and/or (2) the accuracy with which the tongue moves into the closing gesture.

A particularly interesting potential interpretation is that the dips occur where the part of the tongue that is bent behaves most circumferential to the fanlines. Any variation in the tongue contour at the point of constriction is likely to be equidistant from the probe merely shifting perpendicularly to the fanlines rather than varying in distance from the probe. The circumferential shifting along the fanline results in increased variability in that particular area because the tongue contour will be crossing the particular fanline at different slopes for each recording. At the time of consonantal closure, the most convex and also most circumferential part of the tongue is the part that touches the palate. Dips therefore capture the noise in the data that stems from the fact that over multiple repetitions the tongue varies perpendicularly to the fan, with the variation of the most circumferential part (at the most arched part of the tongue) causing the dip.

In our interpretation, slope variation at the most convex part of the tongue contour is the cause for decreased correlation values in the relevant location. Dips indicate where the variation is largest, allowing placement of a vector to measure the kinematics of the stop in the relevant location. Dips may therefore be useful to obtain inform about the articulatory stability of a speaker. The location, steepness and width of the dip may serve as an indicator of how consistently closures are produced across repetitions.

This approach may provide information about coarticulatory processes. The approach is relatively speaker independent though it has its limitations as some speakers' oral cavities or articulation appear to be too far from typical such as , for example, speaker A in the top left panel of Figure 3. Further, measuring the tongue surface movement along the vector that crosses the dip (i.e., the measurement vector) can inform about displacement, velocity and duration of articulatory movement strokes from dynamic data. In contrast to attempts to establish an external referent, dips are inherent to the data, rendering external referents superfluous.

References

Articulate Instruments Ltd (2012). *Articulate Assistant Advanced User Guide: Version 2.14*. Edinburgh, UK: Articulate Instruments Ltd.

Frisch, S. A. (2010). Semi-automatic measurement of stop-consonant articulation using edgetrak. *Ultrafest V*.

Iskarous, K. (2005). Patterns of tongue movement. *Journal of Phonetics*, *33*(4), 363-381.