Tongue movements in space and time: Exploring a dynamic view on speech articulation

Alessandro Vietti

[in collaboration with Alessia Pini*, Simone Vantini* and Lorenzo Spreafico]
Free University of Bozen-Bolzano | *MOX – Politecnico di Milano (Italy)

Research in sociophonetics has prototypically focused on the static analysis of acoustic properties at the midpoints of segments. In recent years, a growing body of evidence contributed to enrich this pointwise analysis in at least two ways showing that:

(1) sociophonetic variation can be found both at the sub-segmental and the supra-segmental level (Docherty & Foulkes 2006), thus indicating that the social meaning is not necessarily attached to the phonetic segment;
(2) the use of articulatory data, especially those obtained from imaging techniques, can expand the domain of analysis, e.g. complementing the acoustic perspective on sound variation and change with information on tongue shapes (Lawson et al. 2013).

In order to contribute to a dynamic view on sociophonetic variation within an articulatory approach, in this talk I introduce a new method for the investigation of tongue shapes data extracted from ultrasound images. The method, named Partial Differential Interval-wise testing (PD-IWT, Pini & Vantini 2015), is based on Functional Data Analysis (Ramsay et al. 2009) and represents an attempt to model and compare ultrasound tongue data in time. PD-IWT helps answering the following research questions:

(a) Are the *n* groups of tongue curves statistically different in space and/or time?
(b) Which regions of the tongue are different?
(c) What is the probability that these differences are occurring by chance?

The partial differential analysis used here offers various metrics for the comparison among groups of tongue profiles. First, as for the spatial dimension, the various tongue profiles are compared among each other as to detect and quantify difference in tongue position, slope and concavity in time. Second, as for the temporal dimension, tongue profiles are compared in terms of velocity and acceleration. Therefore, in this spatiotemporal model of tongue movements, the functional objects are groups of surfaces rather than curves and the output of the analysis is a detailed representation of the statistical differences among groups of surfaces, namely the sequences of sounds under investigation. Those differences are represented as continuous p-value function that specifies the significance of the difference for every region in space and time. As an example, the p-value function is graphically displayed as a heat map in Fig. 1.

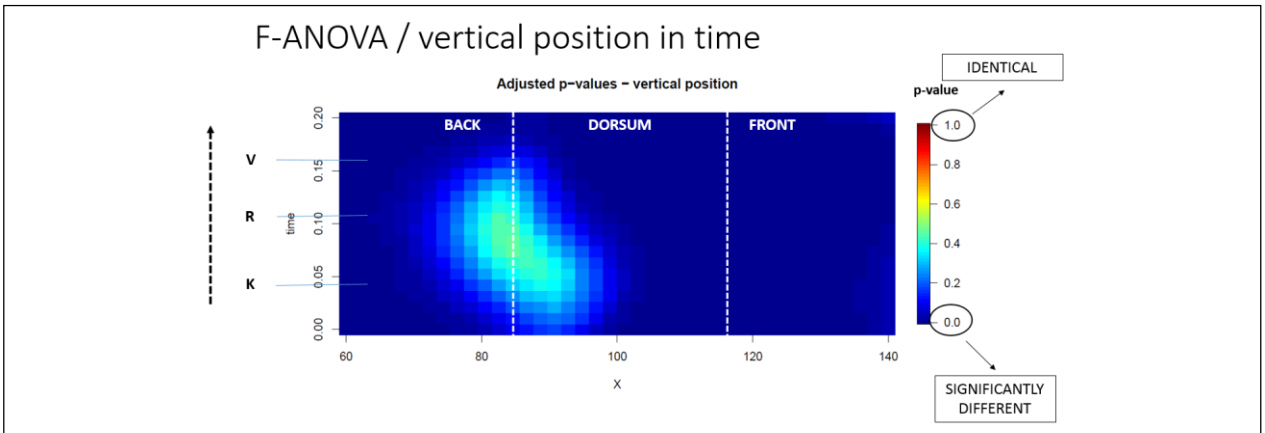

Figure 1 – P-value map globally testing the differences among [kra, kri, kro]: x axis represent the horizontal dimension in the Cartesian plane; on the vertical dimension is represented time (from 0 to 200ms); the α values are represented in deep blue for highly significant difference ($p < .01$) and in light blue, yellow or red for less to non-significant differences ($p > .05$ or $>.10$)

In the presentation, PD-IWT is tested against a corpus of data collected for a study on German-Italian bilinguals in South Tyrol (Italy). The goal is to search for differences in time and/or space in the following two cases: (a) the production of the syllables [kra], [kri], [kro] German; (b) the production of the syllable [tra] in Italian and German by the same bilingual speaker. The outcome is a representation of lingual articulation that captures at the same time patterns of coarticulation, partial independence of the regions of the tongue and the language effect on the spatio-temporal organization of similar segmental sequences.
These results will be discussed in the light of the key theoretical and methodological issues that PD-IWT raises. In particular, I will explore how the analysis of spatiotemporal lingual data fosters a segment-independent understanding of sociophonetic variation. In this view, social experience may be encoded in the kinematic implementation of a word, a sentence or even a larger communicative unit. These units can be understood as emerging categories carved on the basis of variable sociolinguistic experience or driven by the speaker's communicative needs. Transitions between segments, patterns of coarticulation, differential velocities in sub-articulators, sequences of phones, clusters of consonants, syllables structures or any other sort of socio-semiotic valuable units are likely to surface from such a dynamic view of speech production (Hawkins 2003).

References

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics, 34*(4), 409-438.

Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics, 31*(3), 373-405.

Lawson, E., Scobbie, J. M., & Stuart-Smith, J. (2013). Bunched /r/ promotes vowel merger to schwar: An ultrasound tongue imaging study of Scottish sociophonetic variation. *Journal of Phonetics, 41*(3-4), 198-210.

Pini, A. and Vantini, S. (2015), Interval-wise testing for functional data. *Technical Report 30/2015*, MOX – Department of Mathematics, Politecnico di Milano.

Ramsay, J. O., Jooker, G., and Graver, S. (2009). Functional data analysis with R and MATLAB. Springer.