The phonetic basis of the guttural natural class in Levantine Arabic: Evidence from coarticulation and energy components using Deep Learning and Random Forests

Jalal Al-Tamimi (Newcastle University); Emmanuel Ferragne (Paris Diderot University)

Background: Guttural consonants are considered a natural class (McCarthy, 1994) based on evidence that its members phonologically pattern with each other. Membership in this class is not always clear: uvulars and pharyngealized consonants are sometimes included (Sylak-Glassman, 2014; but see McCarthy, 1994). According to the Laryngeal Articulator Model (LAM, Esling, 2005), pharyngeals trigger a maximal epilaryngeal constriction whereas uvulars and to some extent pharyngealized consonants do not. From a coarticulatory point of view, it is also assumed that the guttural class favours regressive, rather than progressive, spreading; and this is even stronger in pharyngealized consonants (Hellmuth, 2013). Our aim is to evaluate the role of coarticulatory patterns and the areas of interest in the signal by looking at how gutturals pattern with each other using two novel approaches. **Method:** 10 Levantine Arabic speakers (5 females), aged 25-45 were recorded producing 21 consonants in Arabic in a /ʔVVCVV/ frame with three subsequent repetitions (VV: symmetric /iː aː uː/; C (6 classes): plain /t d ð s z l/, velar /k ɡ x ɣ/, uvular /q/, pharyngealized /tˤ dˤ ðˤ sˤ zˤ lˤ/, pharyngeal /ħ ʕ/ and glottal /h ʔ/; n = 2034 items). **Analysis:** Spectrograms were fed into a Convolutional Neural Network (CNN) that performed classification and highlighted relevant spectral regions. Next, we used VoiceSauce (Shue et al, 2011) to extract various acoustic metrics (formant, energy and harmonic differences) at multiple data points. Then Principal Component Analyses (PCA) and Random Forests (RF) were used as data reduction technique and predictive tools. **Results:** The CNN achieved a rate of 87% on the two classes (guttural vs non-guttural), and 86% on the six classes. Results presented in Figure 1 (A and B) showed a predominant region of interest within the guttural class (uvular, pharyngeal and pharyngealized) around the low-to-mid frequency range and within the following vowel; the differences between the two groups are located within the first half and the centre of the following vowel. Within the six categories (Figure 1, C to H), the region of interest are common within classes: within gutturals (uvular, pharyngealized and pharyngeal), these are located towards the first half to the middle of the following vowel; with non-gutturals (plain, velar and glottal), these are located towards the middle to end of the preceding vowel to the beginning and middle of the following vowel). The variable frequency locations are indicative of coarticulatory differences between the six categories. To further understand these results, PCA on the first half of the vowel following the consonant showed a clear separation between the 6 categories, however, uvular, pharyngeal and pharyngealized clustered together (Figure 1, I). Glottals are the closest to gutturals but show shared correlates with plains and velars. Our RF results showed a classification rate of 91.3% on the two classes: guttural vs non-guttural and 74% on the six categories; most of the confusions were within the group of gutturals. **Conclusion:** Our results show empirical evidence that is compatible with the legitimacy of the guttural natural class. The regions of interest revealed by the CNN helped us to fine-tune our subsequent analyses by looking specifically at the following vowel in differentiating the two classes, which confirmed that the coarticulatory patterns within gutturals are progressive, with a mixture of patterns within non-gutturals. A combination of formant, energy and harmonic-based metrics provide explanation for this pattern and acoustic evidence that uvular, pharyngeal and pharyngealized consonants induce a gradient epilaryngeal constriction with both laryngeal and supra-laryngeal consequences, contra the categorical predictions of LAM. Gutturals shared use of the energy components around 1-3 kHz that correlated well with the laryngeal and supra-laryngeal changes caused by an epilaryngeal constriction, demonstrating the strength of this combined approach.

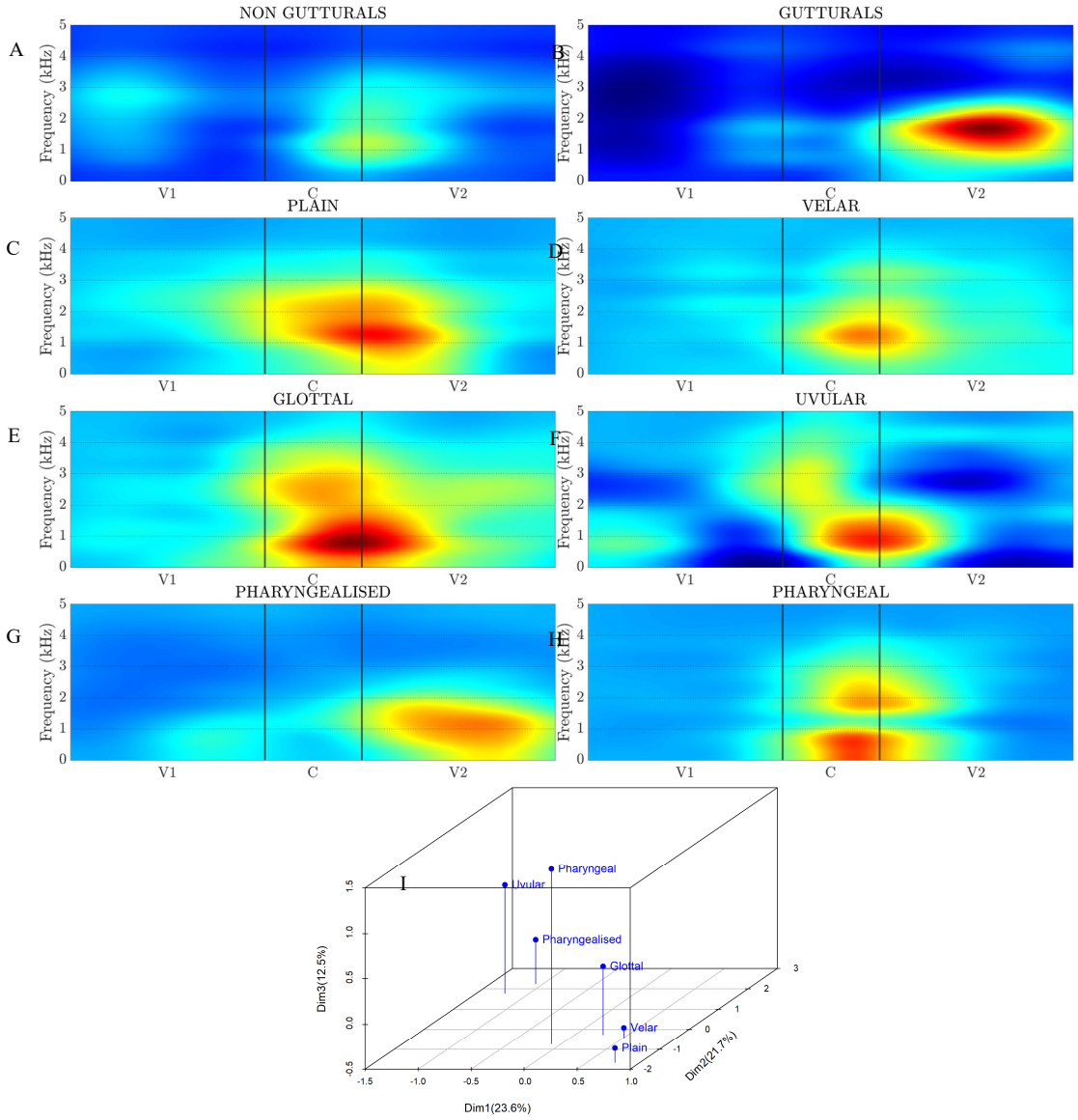

Figure 1: CNN mean activations for non-guttural (A) and Guttural contexts (B) and each of Plain (C), Velar (D), Glottal (E), Uvular (F), Pharyngealized (G), and pharyngeal (H), with blue = least, yellow = partial, red = maximal; 3D PCA of the six categories for V2 Half (I);

References:

Esling, J. (2005). There Are No Back Vowels: The Laryngeal Articulator Model. *The Canadian Journal of Linguistics/La revue canadienne de linguistique*, 50(1), 13–44.

Hellmuth, S. (2013). Phonology. *In: Owens, J. (Ed.), The Oxford handbook of Arabic linguistics*, 45–70. Oxford University Press.

McCarthy, J. J. (1994). The phonetics and phonology of Semitic pharyngeals. *In: Keating, P. (Ed.), Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, 191–233. Cambridge University Press

Shue, Y.-L., P. Keating, C. Vicenik, and K. Hu (2011). VoiceSauce: A program for voice analysis. *In Proceedings ICPhS XVII*: 1846-1849.

Sylak-Glassman, J. (2014). An Emergent Approach to the Guttural Natural Class. *Proceedings of the Annual Meetings on Phonology*, 1–12.