Network analytics reveals patterns in many-to-one articulatory-to-acoustic strategies

Christopher Carignan

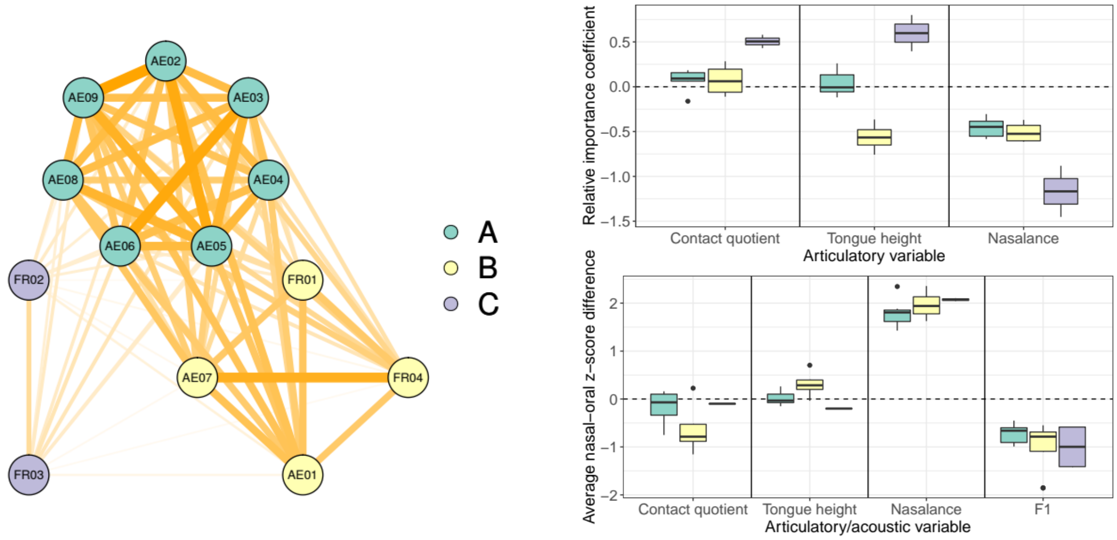
Institute of Phonetics and Speech Processing - LMU Munich

With the many degrees of freedom involved in the articulation of speech sounds, any given acoustic dimension is rarely (if ever) modified by a single articulatory dimension. Rather, the acoustic consequences of speech are often influenced by multiple articulatory dimensions in a many-to-one relationship. A considerably ill-understood relationship of this sort involves the realization of F1 frequency in the production of vowel nasalization. Lowering the velum in order to couple the nasal cavity to the oro-pharyngeal cavity has been shown to modify the frequency of F1 in a manner that is independent of changes in tongue configuration [1]. However, both phonetic and phonological vowel nasality has been observed also to involve changes in tongue configuration [2-4] and increased breathy voicing [5,6], articulations which may modify F1 frequency in their own independent ways [7,8].

Articulatory data (nasalance, ultrasound, EGG) and acoustic data were co-registered to quantify the production of nasalization, tongue height, breathy voicing, and F1 frequency in the distinction of three nasal-oral vowel contrasts in French (FR). These data were collected first from four native FR speakers and, subsequently, from nine naïve Australian English (AE) listeners who imitated the FR productions [9]. For each of the 13 speakers, articulatory measurements were mapped to F1 measurements for each of the nasal-oral vowel pairs using relative importance analysis (RIA), and the RIA coefficients were used to create similarity scores among all of the speakers. These similarity scores were then used to build network models for each vowel pair, and the spinglass algorithm [10] was employed to identify communities/groups of shared articulatory-to-acoustic strategies within each network.

An example of the three-group network for the vowel pair /ɑ̃/-/a/ is shown in Figure 1. The group affiliations for the individual speakers are shown in subfigure (a). Although the overall variable distinctions are similar for the three groups (subfigure (b), bottom panel), the network model is able to identify differences in how these sub-groups of speakers use the three articulators to reach a similar acoustic output (subfigure (b), top panel): using only degree of nasalization (group A), using degree of nasalization and tongue height (group B), or using degree of nasalization, tongue height, and degree of breathiness (group C). In some cases, these strategies are language-dependent (groups A and C), while in other cases similar strategies are used by both FR and AE speakers (group B). The variety of strategies highlights the multi-dimensional nature of vowel nasality, rather than the uni-dimensional assumption of “nasal” vowels as merely oral vowels produced with a lowered velum [11]. Moreover, this multi-dimensionality is not necessarily limited to a native speaker’s phonological system, but it can be transmitted to listener-turned-speakers, which has important implications for models of listener-based sound change that involve the imperfect transmission of articulatory cues between speakers and hearers [12].

This study not only emphasizes the importance of recognizing the many-to-one nature of speech, but it also highlights the necessity of using both acoustic and articulatory data in research on speech production. In the example shown here, community distinctions are not observed in the measurements of the individual articulations, but are only revealed in the articulatory-to-acoustic strategies used by the speakers. Knowledge of these articulatory-to-acoustic relationships could not have been accurately predicted from the articulatory data alone, nor from the acoustic data alone. Research that seeks to gain a more comprehensive understanding of the complex nature of speech production should necessarily involve elements of speech articulation, speech acoustics, and the relationships between the two.



(a) 3-group network for /ɑ̃/-/a/

(b) RIA coefficients and variables

Figure 1. *Spinglass community network results for the /ɑ̃/-/a/ distinction (three groups). The network for the groups is displayed on the left sub-figure (a). The connection strength between any two participant nodes is represented by both line thickness and opacity. Values for each network group are shown in the right sub-figure (b) for the relative importance analysis coefficients (top panel) and the nasal-oral variable distinctions (bottom panel).*

[1] Carignan, C. (2018). Using ultrasound and nasalance to separate oral and nasal contributions to formant frequencies of nasalized vowels. *Journal of the Acoustical Society of America*, *143*(5):2588–2601.

[2] Barlaz, M., Shosted, R., Fu, M., and Sutton, B. (2018). Oropharygneal articulation of phonemic and phonetic nasalization in Brazilian Portuguese. *Journal of Phonetics*, *71*:81–97.

[3] Carignan, C., Shosted, R., Shih, C., and Rong, P. (2011). Compensatory articulation in American English nasalized vowels. *Journal of Phonetics*, *39*:668–682.

[4] Zerling, J. P. (1984). Phénomènes de nasalité et de nasalization vocaliques: Étude cinéradiographique pour deux locuteurs. *Travaux de l'Institut de Phonétique de Strasbourg*, *16*:241–266.

[5] Garellek, M., Ritchart, A., and Kuang, J. (2016). Breathy voice during nasality: A cross-linguistic study. *Journal of Phonetics*, *59*:110–121.

[6] Carignan, C. (2017). Covariation of nasalization, tongue height, and breathiness in the realization of F1 of Southern French nasal vowels. *Journal of Phonetics*, *63*:87–105.

[7] Stevens, K. N. (2000). *Acoustic phonetics*. MIT Press, Cambridge, MA.

[8] Lotto, A. J., Holt, L. L., and Kluender, K. R. (1997). Effect of voice quality on perceived height of English vowels. *Phonetica*, *54*:76–93.

[9] Carignan, C. (2018). Using naïve listener imitations of native speaker productions to investigate mechanisms of listener-based sound change. *Laboratory Phonology*, 9(1):1–31.

[10] Csardi, G. and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, Complex Systems:1695.

[11] Shosted, R. (2015). Nasal vowels are not [+nasal] oral vowels. In Smith, J. and Ihsane, T., editors, *Proceedings of the 42nd Linguistic Symposium on Romance Languages (2012)*, Amsterdam. John Benjamins.

[12] Beddor, P. S. (2009). A coarticulatory path to sound change. *Language*, *85*(4):785–821.