Semantic measures determining coarticulatory movements of the tongue tip

Motoki Saito, Fabian Tomaschek, Harald Baayen
University of Tübingen

According to the tradional view of the speech production process (Levelt, Roelofs, & Meyer, 1999), frequency of occurrence and words' morphological structure should not modulate the fine details of the phonetic realizations of segments. However, a number of studies have observed systematically different coarticulation patterns of phones that depend on word frequency (Tomaschek, Arnold, Bröker, & Baayen, 2018; Tomaschek, Tucker, Fasiolo, & Baayen, 2018; Tomaschek, Wieling, Arnold, & Baayen, 2013). In these studies, morphological status, syllable structure, and frequency were confounded. This makes it difficult to isolate the effect of morphological structure on articulatory movements. The present study investigates tongue movements recorded with electromagnetic articulography (EMA) during stem vowels of German words that share the same stem vowels, i.e. [a(ː)], and word-final segments, i.e. [-t], with/without a morphological boundary between them. The EMA data were extracted from the corpus compiled by Arnold and Tomaschek (2016).

A Generalized Additive Mixed-effects Model (GAM) was fitted to the EMA data, with word frequency and morphological status as predictors of central interest, and with the vertical position of the tongue tip sensor as response variable. This model was pitched against an alternative GAM model with semantic measures derived from a linear discriminative learning (LDL) network (Baayen, Chuang, Shafaei-Bajestan, & Blevins, 2019). One measure, `SupportReceived`, quantified the amount of support received from a word's semantics for a transitional trigram (i.e., a trigram straddling the stem vowel and the word-final segment). A second measure, `SupportGiven` represented the contribution of a transitional trigram to expressing a word's semantics, calibrated against the contribution made by the stem trigram (i.e., trigrams centered on stem vowels). The model with the two learning-based measures outperformed tho model with frequency and morphological status.

Figure 1 shows how tongue contours change over time for different values of `SupportReceived`. The tongue tip is positioned higher in the oral cavity when more suppport for the transitional trigram is received from the semantics. Some anticipatory co-articulation with the final [t] is visible at the right edge of the plot. Figure 2 represents tongue contours along time for `SupportGiven`. Greater values of this predictor indicate that the stem trigram contributes more to expressing the word's meaning than does the boundary trigram. For median values of `SupportGiven`, highlighted by the grey horizontal line, the vertical position of the tongue doesn't change much. For lower and higher values, away from the central value, the tongue starts out at a low position and is raised in anticipation of the final [t]. This suggests that articulatory effort is minimized for the most likely values of `SupportGiven`, and that articulatory effort is invested in initial lowering of the tongue, to produce a more distinct [a], the more unexpected the value of `SupportGiven` is.

The effect of `SupportReceived` is in line with the signal enhancement hypothesis of Kuperman, Pluymaekers, Ernestus, and Baayen (2007). The effect of `SupportGiven` suggests that the signal can not only be enhanced under strong paradigmatic support, but that enhancement can also take place when a triphone contributes either to a very frequent meaning, or to very infrequent meanings. Both effects illustrate that the fine phonetic details of articulation are co-determined by a word's semantics, and therefore challenge traditional models of speech production (Levelt et al., 1999).

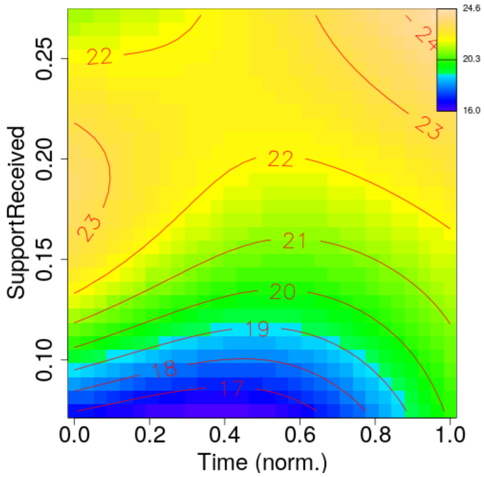

Figure 1: Vertical tongue tip positions as a function of time and `Supportreceived`. Warmer colors represent higher tongue tip positions.

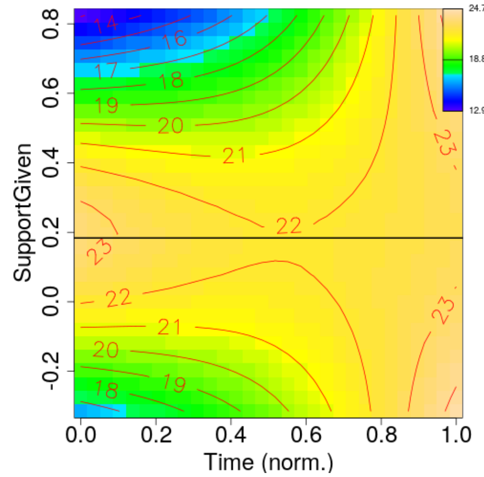

Figure 2: Vertical tongue tip position as a function of time and `SupportGiven`. Higher and lower values of `SupportGiven` characterize words with more informative stems and more informative suffixes respectively.

References

Arnold, D., & Tomaschek, F. (2016). The Karl Eberhards Corpus of spontaneously spoken southern German in dialogues — audio and articulatory recordings. *Tagungsband der 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum*, 9–11.

Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The Discriminative Lexicon: A Unified Computational Model for the Lexicon and Lexical Processing in Comprehension and Production Grounded Not in (De)Composition but in Linear Discriminative Learning. *Complexity*, 1–39. doi: 10.1155/2019/4895891

Kuperman, V., Pluymaekers, M., Ernestus, M., & Baayen, H. (2007). Morphological predictability and acoustic duration of interfixes in Dutch compounds. *The Journal of the Acoustical Society of America*, *121*(4), 2261–2271. doi: 10.1121/1.2537393

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.

Tomaschek, F., Arnold, D., Bröker, F., & Baayen, R. H. (2018). Lexical frequency co-determines the speed-curvature relation in articulation. *Journal of Phonetics*, *68*, 103–116.

Tomaschek, F., Tucker, B. V., Fasiolo, M., & Baayen, R. H. (2018). Practice makes perfect: the consequences of lexical proficiency for articulation. *Linguistics Vanguard*, *4*.

Tomaschek, F., Wieling, M., Arnold, D., & Baayen, H. (2013). Word frequency, vowel length and vowel quality in speech production: An EMA study of the importance of experience. In *Proceedings of the 14th annual conference of the international speech communication association (interspeech 2013)* (pp. 1302–1306).