Introduction: Situating phonological contrast within the perception-production loop

Jason A. Shaw (Yale University)

Some level of correspondence between knowledge of phonological patterns and language-specific speech perception/production is to be expected. Tacit knowledge of these domains co-develop within a speech community. Speech production develops to preserve language-specific systems of contrast; speech perception develops along dimensions of contrast, also on a language-specific basis. Phonological knowledge can be naturally situated as an intermediary of production and perception behavior.

How stringently phonological knowledge of contrast conditions high dimensional speech behaviors is often theory-specific and requires linking hypotheses to bridge theory and data as well as consideration of other interacting factors. This workshop aims to foster discussion of the relation between low dimensional phonological patterning and high dimensional speech behaviors from multiple theoretical perspectives, including Articulatory Phonology and Distinctive Feature Theory. By way of introducing this workshop, I discuss two examples: (1) how gesture coordination as a basis for complex segments makes unique predictions for speech production; (2) how contrastive feature hierarchies can relate to language-specific perceptual behavior.

1. Gestural coordination as a basis for complex segments

Phonological representations in Articulatory Phonology consist of gestures, the atoms of phonological contrast, and coordination relations between them (Gafos & Goldstein, 2012). Patterns of gesture coordination have been found to correspond to aspects of phonological structure, such as syllables (e.g., Hermes, Mücke, & Grice, 2013; Krakow, 1999; Shaw & Gafos, 2015), and to relate systematically to phonological processes, e.g., vowel harmony (Benus & Gafos, 2007), allophony (Browman & Goldstein, 1990), morphological template filling (Gafos, 2002). Expressing phonological representations as patterns of gestural coordination commits to nuanced phonetic consequences, which can be tested in kinematic data or, in some cases, with the resulting acoustics. For example, Shaw et al. (2019) hypothesize that gestural coordination is what differentiates complex segments, such as palatalized consonants in Russian, /bʲ/, /mʲ/, /vʲ/, from similar consonant sequences /bj/, /mj/, /vj/. Specifically, the hypothesis is that gestures of complex segments are coordinated with reference only

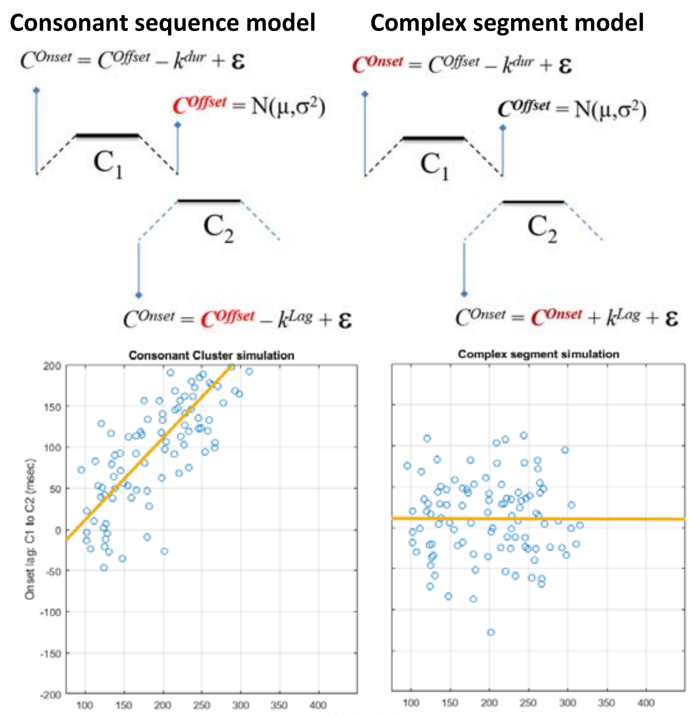

Figure 1: (top) stochastic models of gesture coordination for consonant clusters (left) and complex segments (right); (bottom) simulation results for each model showing predicted relation between C1 duration (x-axis) and onset lag (y-axis)

to the temporal onsets of the gestures while gestures of consonant sequences are coordinated with reference to the temporal offset of the first gesture (C_1) and the onset of the second (C_2).

Figure 1 shows stochastic models based on these coordination relations (top) and simulated kinematic measures (bottom). Crucially, in this formulation it is not the degree of gesture overlap which is related to phonological structure but the temporal dependency between articulatory landmarks. The simulation results reveal that complex segments and clusters differ in patterns of covariation between C_1 DURATION (x-axis) and ONSET LAG (y-axis), the temporal lag between gesture onsets. For consonant clusters (left), an increase in C_1 gesture duration conditions a corresponding increase in onset lag. For complex segments (right), variation in C_1 duration has no effect on onset lag. *Thus, different patterns of gestural coordination impose distinct patterns of covariation between temporal intervals*. For the case considered here, these predictions correspond closely to the data: Russian palatalized consonants show the complex segment pattern (Figure 1: right) while English consonant-glide sequences show the consonant cluster pattern (Figure 1: left)(Shaw, Durvasula, & Kochetov, 2019). More broadly, languages differ in whether sounds that are articulatorily complex behave in the phonology as sequences of segments or as single complex segments. In some languages, clear evidence comes from phonological patterning, including distributional evidence; in others, such evidence is lacking. If the phonological patterning is linked systematically to coordination, then language-specific speech production goes hand in hand with acquiring the appropriate phonological structure (complex segment vs. cluster). Representing phonological structure as characteristic patterns of gestural coordination (Browman & Goldstein, 1988) commits to nuanced predictions for speech production.

2. Contrastive feature hierarchies as a basis for language-specific perception

Contrastive Feature Hierarchies constructed via the successive division algorithm (Dresher, 2009) predict that similar (or identical) segmental inventories can have different feature structure. Consider, for example, the partial feature hierarchies for vowels in Australian vs. New Zealand English (Figure 2: top). The accents have similar numbers of vowels but the vowels differ across accents in their phonetic realizations. On the hypothesis that listeners parse ambient speech according to their phonological system (e.g., Hwang, Monahan, & Idsardi, 2010; Poeppel, Idsardi, & van Wassenhove, 2008), differences in feature hierarchies may drive different patterns of perceptual confusion across accents. Vowel categorization experiments, analyzed via hierarchical cluster analysis, reveal accent-specific patterns of perceptual confusion (Shaw, Best, et al., 2019). Listeners across accents attune to phonetic dimensions in different ways, which feature hierarchies have the potential to explain.

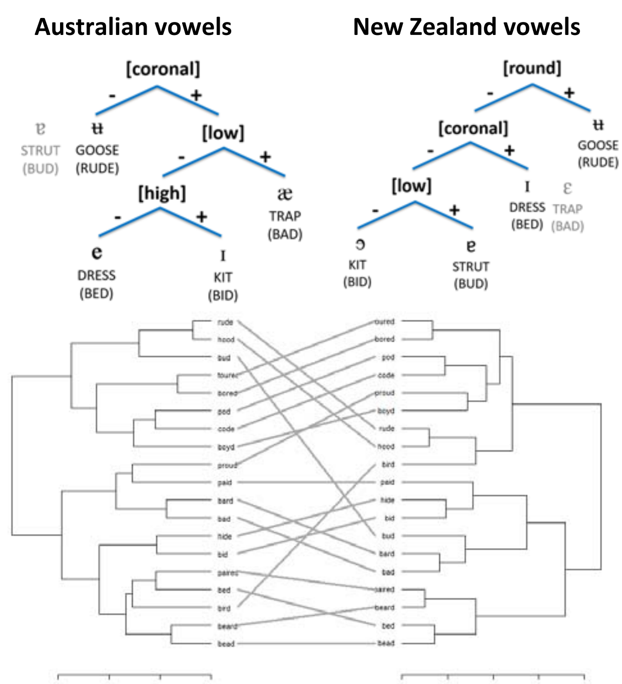

Figure 2: (top) partial feature hierarchies; (bottom) tanglegram comparing perceptual confusions

Introduction: Situating phonological contrast within the perception-production loop

The tanglegram in Figure 2 (bottom) compares perceptual similarity of Australian (left) and New Zealand (right) vowels. The nodes are reference words selected by listeners to classify vowel stimuli produced in nonce words. The first branch of the dendrogram for Australian listeners splits coronal vowels from non-coronal vowels; the first branch in the New Zealand dendrogram instead splits round vowels from unrounded vowels. Evidence that accent-specific perception patterns are stable, in a way expected of phonological structure, comes from the cross-accent perceptual scenario—listeners tend to impose the perceptual structure of their native accent on the vowels of other accents (Shaw et al., 2018), a pattern which generalizes across multiple accent pairs (Shaw et al., 2019).

References

Benus, S., & Gafos, A. I. (2007). Articulatory characteristics of Hungarian 'transparent'vowels. *Journal of Phonetics, 35*(3), 271-300.

Browman, C., & Goldstein, L. (1990). Tiers in Articulatory Phonology with some implications for casual speech. In Kingston & Beckman (Eds.), *Papers in Laboratory Phonology I: between the grammar and physics of speech* (pp. 341-376). Cambridge: Cambridge University Press.

Browman, C. P., & Goldstein, L. (1988). Some Notes on Syllable Structure in Articulatory Phonology. *Phonetica, 45*, 140-155.

Dresher, B. E. (2009). *The contrastive hierarchy in phonology* (Vol. 121): Cambridge University Press.

Gafos, A. (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory, 20*, 269-337.

Gafos, A., & Goldstein, L. (2012). Articulatory representation and organization. In A. C. Cohn, C. Fougeron, & M. K. Huffman (Eds.), *The Oxford Handbook of Laboratory Phonology* (pp. 220-231).

Hermes, A., Mücke, D., & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: The case of "impure s". *Phonology, 30*(1), 1-25.

Hwang, S.-O. K., Monahan, P. J., & Idsardi, W. J. (2010). Underspecification and asymmetries in voicing perception. *Phonology, 27*(2), 205-224.

Krakow, R. A. (1999). Physiological Organization of Syllables: A Review. *Journal of Phonetics, 27*(1), 23-54.

Poeppel, D., Idsardi, W., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos Trans R Soc Land B, 363*, 1071-1086.

Shaw, J. A., Best, C., Docherty, G., Evans, B. G., Foulkes, P., Hay, J., & Mulak, K. E. (2018). Resilience of English vowel perception across regional accent variation. *Laboratory Phonology, 9*(1), 1-36. doi: http://doi.org/10.5334/labphon.87

Shaw, J. A., Best, C. T., Docherty, G., Evans, B., Foulkes, P., Hay, J., & Mulak, K. (2019). An information theoretic perspective on perceptual structure: cross-accent vowel perception. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.) *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019 (pp. 582-586). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Shaw, J. A., Durvasula, K., & Kochetov, A. (2019). The temporal basis of complex segments. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.) *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019 (pp. 676-680). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Shaw, J. A., & Gafos, A. I. (2015). Stochastic Time Models of Syllable Structure. *PLoS One, 10*(5), e0124714 0124711-0124736.