Editorial: models and theories of speech production

Health & Medicine

Editorial on the Research Topic

Models and Theories of Speech Production

Spoken language is conveyed via well-coordinated speech movements, which act as coherent units of control referred to as gestures. These gestures and their underlying movements show several distinctive properties in terms of lawful relations among the parameters of duration, relative timing, range of motion, target accuracy, and speed. However, currently, no existing theory successfully accounts for all properties of these movements. Even though models in speech motor control in the last 40 years have consistently taken inspiration from general movement science, some of the comparisons remain ill-informed. For example, our present knowledge on whether widely known principles that apply to limb movements (e. g., the speed-accuracy trade off known as Fitts' law) also hold true for speech movements is still very limited. An understanding of the principles that apply to speech movements is key to defining the somewhat elusive concept of speech motor skill and to assessing and interpreting different levels of that skill in populations with and without diagnosed speech disorders. The latter issue taps into fundamental debates about whether speech pathology assessment paradigms need to be restricted to control regimes that are specific to those underlying typical speech productions. Resolution of such debates crucially relies on our understanding of the nature of speech processes and the underlying control units.

Unlike movements in locomotion or oculomotor function, speech movements when combined into gestures are not mere physical instantiations of organs moving in space and time but, also, have intrinsic symbolic function.

Language-particular systems, or phonological grammars, are involved in the patterning of these gestures. Grammar constraints regulate the permissible symbolic combinations as evidenced via eliciting judgments on whether any given sequence is well-formed in any particular language (the same sequence can be acceptable in one, but not the other language). In what ways these constraints shape speech gestures and how these fit with existing general principles of motor control is, also, not clearly understood.

Furthermore, speech gestures are parts of words and thus one window into understanding the nature of the speech production 1 system is to observe speech movements as parts of words or larger chunks of speech such as phrases or sentences. The intention to produce a lexical item involves activating sequences of gestures that are part of the lexical item. The regulation in time of the units in such sequences raises major questions for speech motor control theories (but also for theories of cognition and sequential action in general). Major challenges are met in the inter-dependence among different time scales related to gestural planning, movement execution and coordination within and across domains of individual lexical items. How these different time scales interact and how their interaction affects the observed movement properties is for the most part still unknown.

In this special issue, we present a variety of theoretical and empirical contributions which explore the nature of the dynamics of speech motor control. For practical purposes, we separate these contributions in two major themes:

1) Models and theories of speech production.

2) Applications.

Following is a short description of each paper as listed under these themes.

1) Models and theories of speech production

The speech signal is simultaneously expressed in two information-encoding systems: articulation and acoustics. Goldstein's contribution addresses the relation between representations in these two parallel manifestations of speech while focusing not on static properties but on patterns of change over time (temporal co-modulation) in these two channels. To do so, Goldstein quantifies the relation between rates of change in the parallel acoustic and articulatory representations of the same utterance, produced by various speakers, based on x-ray microbeam data. Analysis of this relation indicates that the two representations are correlated via a pulse-like modulation structure, with local correlations being stronger than global ones. This modulation seems linked to the fundamental unit of the syllable.

It is widely assumed that acoustic parameters for vowels are normally distributed, but it is rarely demonstrated that this might be the case. Whalen and Chen quantified the distributions of F1 and F2 values of /i/ and /o/ in the English words “ heed,” “ geek,” “ ode”/“ owed,” and “ dote” produced by a single speaker on three different days. Analysis based on a high number of repetitions of these vowels in different consonantal contexts indicates that distributions are generally normal, which in turn suggests consistent vowel-specific targets across different contextual environments. The results add

weight to the widely-held assumption that speech targets follow a normal distribution and the authors discuss the implications for theories of speech targets.

Laganaro, M. (2019). Phonetic encoding in utterance production: a review of open issues from 1989 to 2018. *Language Cognit. Neurosci.* 34, 1193–1201. doi: 10. 1080/23273798. 2019. 1599128

Meyer, A. S., Ardi, R., and Laurel, B. (2019). Thirty years of speaking: an introduction to the Special Issue. *Language Cognit. Neurosci.* 34, 1073–1084. doi: 10. 1080/23273798. 2019. 1652763

Saltzman, E. L., and Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology* , 1, 333–82. doi: 10. 1207/s15326969eco0104_2

Schaal, S., Mohajerian, P., Ijspeert, A. J., Cisek, P., Drew, T., and Kalaska, J. F. (2007). Dynamics systems vs. Optimal control a unifying view. *In Progress in Brain Research 165* , 425–45. doi: 10. 1016/S0079-6123(06)65027-9