QUANTITATIVE ASSESSMENT OF A NEUROCOMPUTATIONAL MODEL OF SPEECH PRODUCTION WITH NEUROIMAGING DATA

Frank H. Guenther[1], Ayoub Daliri[2,1], Alfonso Nieto-Castanon[1], and Jason A. Tourville[1]

[1]Boston University, Boston, MA, 02215, USA

[2]Arizona State University, Tempe, AZ, 85287, USA

Advances in neuroimaging techniques such as functional magnetic resonance imaging (fMRI) and electrocorticography (ECoG) over the past two decades have resulted in a greatly improved understanding of the neural mechanisms underlying human sensory, motor, and cognitive capabilities, leading to increasingly sophisticated neural models of these functions. Within the domain of speech, we have developed and refined a large-scale neurocomputational model, called the Directions into Velocities of Articulators (DIVA) model (Guenther, 2016; Guenther, Ghosh, & Tourville, 2006; Golfinopoulos, Tourville, & Guenther, 2010), which provides a unified mechanistic account of acoustic, kinematic, and neuroimaging data on speech. Functional neuroimaging has been a powerful means for evaluating and refining such models. To date, however, these evaluations have been almost exclusively qualitative. Quantitative evaluations have been hampered by the absence of a general computational framework for (i) generating predicted functional activation from a model that can be directly and quantitatively compared to empirical functional neuroimaging data, and (ii) testing between models to identify the model that best fits experimental data.

Here we present a general computational framework to overcome these issues, with a specific application to the DIVA model. Within this framework, the brain network responsible for a task is broken into a set of computational nodes, each of which is localized to an MNI stereotactic coordinate in the brain. Associated with each node is a computational load function that links the node's activity to a computation involving quantifiable measures from the task. The instantaneous neural activity at each location in the brain (e.g., each voxel of an fMRI image, or each electrode of an ECoG array) is then calculated by summing the contributions of all model nodes at that location, with each node treated as a Gaussian activity source centered at the node's location. The parameters of the Gaussians (i.e., spread and magnitude of activation) are optimized to produce the best fit to the functional data. Model comparisons are based on the overall fit level and number of free parameters using the Akaike Information Criterion (AIC).

This framework was used in conjunction with a large fMRI database of speech production studies (116 speakers) to illustrate the DIVA model's ability to provide a unified account for whole-brain activity patterns seen during speech production under normal and perturbed conditions. [Supported by NIH grants R01 DC002852, R01 DC007683.]

References

Guenther, F. H. (2016). *Neural control of speech*. Cambridge, MA: MIT Press.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language,* 96, 80-301.

Golfinopoulos, E., Tourville, J. A., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *NeuroImage,* 52, 862-874.