Developing a model of speech production using the Neural Engineering Framework (NEF) and the Semantic Pointer Architecture (SPA)

Bernd J. Kröger[1], Trevor Bekolay[2,3], Peter Blouw[2,3], Terence C. Stewart[4]

[1]*Department for Phoniatrics, Pedaudiology and Communication Disorders, Medical Faculty RWTH Aachen University, Germany*

[2]*Applied Brain Research, Waterloo, ON, Canada*

[3]*Centre for Theoretical Neuroscience, University of Waterloo, Waterloo, ON, Canada*

[4]*National Research Council of Canada, University of Waterloo Collaboration Centre, Waterloo, ON, Canada*

Abstract

Only a few approaches exist that are capable of combining the symbolic linguistic part of speech production with the phonetic sensory-motor part. The latter is especially difficult to model due to the need for a concrete articulatory-acoustic model that is controlled by a motor module activating a temporally well synchronized set of speech articulator movement units (SAMUs) based on the output of the phonological-linguistic module [1]. In addition, a stream of realistic somatosensory (tactile as well as proprioceptive) and auditive feedback signals is needed. The generation of realistic feedback signals requires a high-level articulatory-acoustic model of an artificial vocal tract as part of the whole production model [1].

In this conference contribution we will explain how a comprehensive symbolic-linguistic and sensory-motor model of speech production can be established using the well-established Neural Engineering Framework (NEF, [2, 3]) for designing large scale neural models [3, 4]. This approach profits from its combination with the Semantic Pointer Architecture (SPA, [3, 4]), which enables us to define and to process neural signals representing symbolic cognitive linguistic units like words, lemmas, and phonological forms of syllables as well as sensory-motor units like SAMUs [1].

Our biologically inspired neural speech production model has already been applied to medical research questions [5, 6, 7, 8] as well as to basic linguistic research questions concerning the feedback mechanisms involved in the production and repair of word production errors [9] but is expanded now by a motor module as well as by lower level auditory and somatosensory feedback mechanisms.

In our talk we will introduce specifically this new approach for generating motor control commands based on a NEF-SPA neural oscillator model [10], which allows us to elegantly generate the set of temporally synchronized SAMUs needed for the production of a syllable or word. A main benefit of this control approach lies in the fact that the speaking rate can be controlled by changing one parameter, i.e., the frequency of the neural syllable generation oscillators. The temporal coordination of all SAMUs is controlled by NEF-SPA phasing rules that define the points in time for starting and ending the activation of lower level neural oscillators controlling single SAMUs. First simulation results will be presented.

References

[1] Kröger BJ, and Bekolay T (2019) Neural Modeling of Speech Processing and Speech Learning. An Introduction. Springer International Publishing. ISBN 978-3-030-15852-1

[2] Eliasmith C, and Anderson CH (2003) Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems. Cambridge, MA: MIT Press

[3] Eliasmith C (2013) How to Build a Brain: A Neural Architecture for Biological Cognition. New York, NY: Oxford University Press

[4] Eliasmith C, Stewart TC, Choo X, Bekolay T, DeWolf T, Tang Y, et al. (2012) A large-scale model of the functioning brain. Science 338, 1202–1205. doi: 10.1126/science.1225266

[5] Stille C, Bekolay T, Blouw P, Kröger BJ (2020) Modeling the Mental Lexicon as Part of Long-Term and Working Memory and Simulating Lexical Access in a Naming Task Including Semantic and Phonological Cues. Frontiers in Psychology, doi: 10.3389/fpsyg.2020.01594

[6] Stille C, Bekolay T, Blouw P, Kröger BJ (2019) Natural language processing in large-scale neural models for medical screenings. Frontiers in Robotics and AI (Computational Intelligence) doi: 10.3389/frobt.2019.00062

[7] Senft V, Stewart TC, Bekolay T, Eliasmith C, Kröger BJ (2018) Inhibiting Basal Ganglia Regions Reduces Syllable Sequencing Errors in Parkinson's Disease: A Computer Simulation Study. Frontiers in Computational Neuroscience 12:41. doi: 10.3389/fncom.2018.00041

[8] Senft V, Stewart TC, Bekolay T, Eliasmith C, Kröger BJ (2016) Reduction of dopamine in basal ganglia and its effects on syllable sequencing in speech: A computer simulation study. Basal Ganglia 6: 7-17

[9] Kröger BJ, Crawford E, Bekolay T, Eliasmith C (2016) Modeling interactions between speech production and perception: speech error detection at semantic and phonological levels and the inner speech loop. Frontiers in Computational Neuroscience 10:51. doi: 10.3389/fncom.2016.00051

[10] Kröger BJ, Bekolay T, Blouw P (2016) Modeling motor planning in speech processing using the Neural Engineering Framework. In: Jokisch O (Ed.) Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2016 (TUDpress, Dresden, Germany), pp. 15-22 (ISBN: 978-3-95908-040-8)