Neural Tracking Measures of Speech Intelligibility: Manipulating Intelligibility while Keeping Acoustics Unchanged

I.M Dushyanthi Karunathilake[1], Joshua P. Kulasingham[2], Jonathan Z. Simon[1,3,4]

[1]*Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA*
[2]*Department of Electrical Engineering, Linköping University, Linköping, SE*
[3]*Institute for Systems Research, University of Maryland, College Park, MD 20742, USA*
[4]*Department of Biology, University of Maryland, College Park, MD 20742, USA*

Neural speech tracking has advanced our understanding of how our brains rapidly map an acoustic speech signal onto linguistic representations and ultimately meaning. However, it remains unclear which aspects of the corresponding neural responses correspond to speech intelligibility, which is only loosely coupled to the acoustics. Intelligibility related neuro-markers derived from such neural responses would play a crucial role in advancing our understanding of the neurophysiology of the speech understanding, evaluation of auditory function across diverse clinical populations, and hearing device evaluation. Many studies addressing this question vary the level of intelligibility by manipulating the acoustic waveform, making it difficult to cleanly distinguish effects of intelligibility from the underlying acoustical confounds. In this study, speech intelligibility is manipulated while keeping the acoustical structure unchanged, using degraded speech plus a priming paradigm. Acoustically identical three-band noise vocoded (degraded) speech segments (~20 s duration) are presented twice, but the second presentation is preceded by the original (non-degraded) version of the same speech segment. This priming, which generates a 'pop-out' percept, substantially improves the intelligibility of the second presentation of the degraded speech passage while keeping the acoustics identical. We recorded magnetoencephalography (MEG) data from 25 younger adults and investigated how intelligibility affects auditory and linguistic neural tracking measures using multivariate Temporal Response Functions (mTRFs). As expected, behavioral results confirmed that perceived speech clarity is improved by priming. mTRF analysis revealed that auditory (speech envelope and envelope onset) and phoneme onset neural responses are influenced only by the acoustics of the sensory input (bottom-up driven mechanisms). Critically, our key findings suggest that neural measures associated with the segmentation of sounds into words emerges first with better speech intelligibility, especially those time-locked at N400-like latencies in prefrontal cortex (PFC), in line with engagement of top-down mechanisms associated with priming. Taken together, our results suggest that time locked neural responses associated with lexical segmentation may serve as novel objective measures of speech intelligibility.

This work was supported by the National Institutes of Health grant R01-DC019394 and National Science Foundation grant SMA 1734892