Hierarchical neural networks for degraded speech processing

Bradley White
Brain and Language Center for Neuroimaging
Gallaudet University

Advances and Perspectives in Auditory Neuroscience 2023
Washington, DC USA

Decades of research have implicated a wide range of cortical areas involved with degraded speech processing and listening effort, but how these areas are functionally organized to perform such complex tasks is not well understood. We investigated how degraded speech impacts neural networks between attention and language areas in the human brain. Optical fNIRS brain imaging data were analyzed from 29 young adults (18 females; mean age = 30.66 ± 6.12). Participants were right-handed, screened for no hearing loss, and assessed for comparable speech recognition, English language understanding, and nonverbal intelligence. Participants were monolingual native English speakers with little-to-no experience with another language. Participants listened to sentences and performed a plausibility judgment task. Sentences were blocked and randomized in a 2x2x3 design that varied in syntax (simple subject-object order and complex object-subject order), speed (typical and fast), and clarity (clear undegraded, 3 channel band-pass filtered, and 8 channel noise vocoded). We predicted that cognitive load from increasingly degraded speech would trade off with executive resources for attention in the prefrontal cortex (PFC) and language in the left hemisphere (LH). Data were preprocessed and analyzed using the NIRS Brain AnalyzIR Toolbox. Subject-level functional connectivity (FC) was modeled between all possible channel pairs for each condition using robust general linear regression with pre-whitening. Group-level FC was modeled per channel pair for each condition using robust linear mixed-effects. Each subject was treated as a random effect. FC between conditions were compared using t- tests and corrected for multiple comparisons. For significant FC in each contrast, a weighted graph model was constructed with nodal degree and PageRank. We observed three main findings. First, band-pass filtered and noise vocoded speech differently impacted FC for syntax. Second, these FC networks were sensitive to interactions with additional listening challenges (speech rate). Third, the PFC had dual roles. During low cognitive demand and disengagement, the PFC operated independently of the LH (no FC). When listening was more degraded, the PFC operated in synchrony with the LH (FC). These findings inform us of the increased computational demands that may be required for successful processing of degraded speech, and how the PFC and LH network together to overcome listening challenges.