Cambridge Language Sciences Annual Symposium 2021: Language and Inclusion, 23 November 2021

Keynote: 'Re-understanding speech understanding: Closing the cohort loop'

William D Marslen-Wilson, Centre for Speech, Language, and the Brain, Department of Psychology, University of Cambridge

Chair: Mirjana Bozic

Human listeners understand spoken language literally as they hear it, reflecting a perceptually seamless process of real-time comprehension of what the speaker is saying. This remarkable experience of immediacy is rooted in the exceptional earliness with which information carried by successive words is integrated into the interpretation of the current utterance. But despite 50 years of research, there has been no accepted mechanistic neurobiological account of the brain systems that support this process. Only recently have scientific tools emerged that allow us to probe the real-time activity of these brain systems, telling us where and when such activity can be detected and what their neurocomputational content might be. The resulting research enables us, first, to reject the historically dominant account of early speech interpretation as a linguistically stratified computational hierarchy, centered around the notion of a phoneme, and based on sequential transitions between successive representational states.

We propose instead a fully distributed non-hierarchical recurrent neural architecture, active across bilateral temporo-parietal cortex, where there is no representational specificity early in the process; simply an optimised neural pathway towards the listener's semantically interpreted speech percept. Second, we have identified a discrete left hemisphere fronto-temporal component that mediates the early integration of bottom-up speech constraints with top-down contextual constraints. Saliently, the time-locked phonological input to this integration process appears to be continuous, already present at (and before) word onset, and likely to be tracking the changing articulatory states of the source. This continuous flow of phonological constraint, providing partial cues to word identity, interacts with partial contextual semantic cues to mediate initial access to lexical form and meaning within 100 ms of word onset in sentential context, supporting an optimally efficient uptake of available cues to speech interpretation, and where the contribution of this early integration component terminates as each word is recognised. These proposals bring the core claims of the original cohort model into an explanatory computational cognitive neuroscience framework.

READ MORE: William Marslen-Wilson speaker profile