Contribution: Poster presentation at the XXIX International Evoked Response Audiometry Study Group (IERASG-2025) Biennial Symposium, Boulder, Colorado, EEUU (June 14-18, 2025)

Title: Understanding early and late transient responses to speech: experiments using synthetic speech.

Authors: Ángel de la Torre[1], Isaac M. Alvarez[1], Nicolás Müller[2], Francisco Chiquero[2], Juan Martín-Lagos[2], José L. Vargas[2]

(1) Department of Signal Theory, Telematics and Communications. CITIC-UGR. University of Granada, Spain
(2) Department of Otorhinolaryngology, Hospital Clínico Universitario San Cecilio, Granada, Spain

Topics: Speech perception/discrimination, ABR; MLR/ALR

Keywords: Multi-response deconvolution, auditory evoked potential (AEP), auditory brainstem response (ABR), middle latency response (MLR), auditory late response (ALR), glottal pulses.

ABSTRACT:

Background: Recording auditory responses evoked with natural speech is challenging, due to the inherent variability and typical parameters of normal speech, with fundamental frequency between 90-260 Hz (glottal pulses separated between 3.8 and 11 ms) and articulation rates around 10-12 phonemes/s (phoneme duration between 40 and 150 ms). Due to this difficulty, conventional methods for obtaining evoked responses to speech include transient responses using short and isolated vowels or syllables, frequency following response to fundamental frequency in stationary phonemes, or speech-evoked envelope following responses.

Methods: Based on the multi-response deconvolution method developed by our team, and using synthetic phonemes, we propose a simultaneous deconvolution of events corresponding to phonemes and to events corresponding to glottal pulses. In the proposed protocol, the stimulation pattern includes several phonemes (/a/, /i/ and /s/) and click bursts. We deconvolve separately the responses to the short-term events (glottal pulses in vowels and individual clicks in click bursts), and to the long-term events (phonemes or click bursts).

Results: Responses recorded from 6 subjects, reveals consistent early responses (with clear ABR/MLR components) to the short-term events and late responses (with clear MLR/ALR components) to the long-term events. Results with different configurations of the stimulation pattern also illustrate which configurations are more appropriate to obtain consistent responses within exploration times around 10 minutes.

Conclusions: This preliminary study with synthetic speech provides valuable information for a better understanding of the perception of the speech, and for the development of protocols based on speech signals to assess the auditory function.