2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Neural representation of physical and perceived environmental acoustics

García-Lázaro, Haydée G. & Teng, Santani
The Smith-Kettlewell Eye Research Institute

***Abstract*— We recorded EEG from healthy human listeners (N=14) as they distinguished real from synthetic reverberant impulse responses (IRs) convolved with speech samples. Neural responses distinguished between real and synthetic IRs starting about 500 ms post-stimulus onset, and predicted subject reports after ~800ms. Our results indicate dissociable neural mechanisms between sensory and higher cognitive processing during auditory scene analysis.**

***Clinical Relevance*— Most real-world hearing takes place in acoustically cluttered, reverberant environments, which makes perceptual segregation of sound sources critical. Reverberation also carries environmental spatial information of potential use to blind and visually impaired persons. Understanding the neural mechanisms of auditory scene analysis can help identify points of failure in high-level hearing loss and guide behavioral or technological therapeutic interventions.**

I. Introduction

Acoustic reverberation, the aggregate sound reflections from multiple nearby surfaces, both distorts source signals (e.g. speech) and carries behaviorally useful information about local space, especially for blind and visually impaired persons. While human listeners can perceive the statistics of reverberant acoustics that facilitate perceptual [1] and neural [2] segregation of direct and reverberant sounds, the direct link between neural representations of reverberant acoustics and their perception remains unclear.

II. Methods

We conducted an electroencephalographic (EEG) study in which sighted participants (N=14) listened to 2s excerpts of spoken sentences, unique to each trial and convolved with a reverberant IR. Each IR was either "Real" (R), recorded in a real environment, or "Fake" (F), synthesized to emulate or deviate from real reverberation [1] (**Fig. 1**). Subjects judged the reverberant space as R or F. We applied multivariate pattern analysis to distinguish R vs. F IRs and predict the perceptual reports from neural responses (**Fig. 2A**).

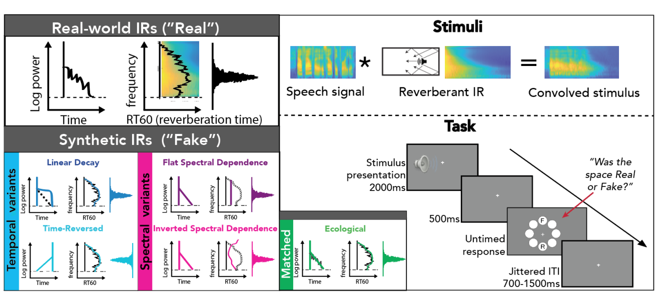

Figure 1. Overview of stimuli and task. Log-power/frequency vs. time diagrams characterize real IRs and their synthesized variants (left). The convolved stimulus (top right) required participants to segregate the reverberation from speech to perform the R vs. F classification task (right).

III. Results

We found reliable classification between R and F IRs starting at ~500 ms. Classifier performance increased when filtering for correct responses, suggesting that neural representations reflect processes underlying sensory encoding as well as task performance (**Fig. 2B**). Likewise, decoding trials by perceptual report improved decoding accuracy over the stimulus condition decoding after ~800 ms (**Fig. 2C**). This suggests that early neural responses to the stimuli were similar between R and F conditions, began to respond to distinguishing features at ~500ms, and transitioned to representing the perceived authenticity of the reverberant background approximately 300 ms later.

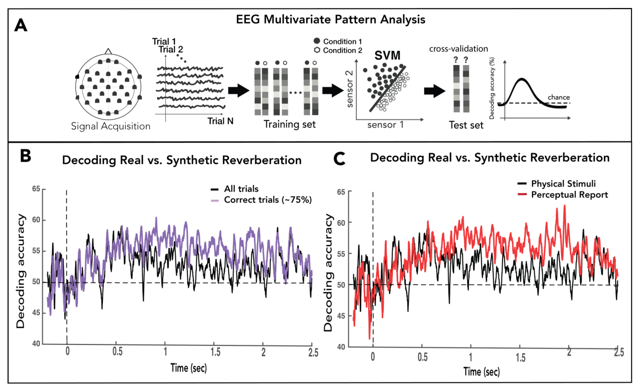

Figure 2. Overview of analysis and decoding results. We applied time-resolved SVM decoding (A) to EEG signals to decode real (R) vs. synthetic ("fake," F) reverberation over the 2s of stimulus presentation up to the onset of the response cue at 2500ms. Effect of filtering for successful trials is shown in B; perceived vs. physical decoding shown in C.

IV. Discussion & Conclusions

The neural response reliably distinguishes subtle statistical regularities of physically realistic reverberation from deviations in multiple dimensions. This indicates a finely tuned mechanism for identifying background sounds and segregating them from direct acoustic signals, with a temporal lag between acoustic encoding and decision processes that predict the eventual response.

Acknowledgments

We thank J. Traer, A. Wong-Kee-You, and Y. Alwis for assistance and commentary in preparing this study. Funding sources: E. Matilda Ziegler Foundation for the Blind; Smith-Kettlewell Eye Research Institute.

References

[1] J. Traer and J. H. McDermott, "Statistics of natural reverberation enable perceptual separation of sound and space," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 48, pp. E7856–E7865, Nov. 2016.

[2] S. Teng, V. R. Sommer, D. Pantazis, and A. Oliva, "Hearing Scenes: A Neuromagnetic Signature of Auditory Source and Reverberant Space Separation," *eNeuro*, vol. 4, no. 1, Jan. 2017, doi: 10.1523/ENEURO.0007-17.2017.