Neural Correlates of Individual Differences in Categorical Perception

Sara D. Beach[1,2], Tracy M. Centanni[2], Ola Ozernov-Palchik[2,3], Sidney C. May[2], Dimitrios Pantazis[2], Tyler K. Perrachione[4], John D. E. Gabrieli[2]

[1]Harvard University; [2]Massachusetts Institute of Technology; [3]Tufts University; [4]Boston University

Speech is a continuous acoustic signal usually perceived categorically. Behaviorally, this looks like:
- less consistent labeling of the same sound over multiple trials (Is this a /ba/ or a /da/?);
- a shallower labeling function slope over an acoustic continuum (/ba/ to /da/ in 10 equal steps);
- weaker discrimination of sounds across a phonemic boundary (Same or different? /ba/ vs. /da/);
- stronger discrimination of sounds within a phonemic category (Same or different? $/ba/_1$ vs. $/ba/_2$).

Atypical, more continuous perception of speech – especially consonants – has been linked to communication/learning disorders such as dyslexia [1]. To understand the cause of this behavioral difference, we investigated the structure of phonetic and phonemic representations in the brain using univariate and multivariate analyses of MEG responses.

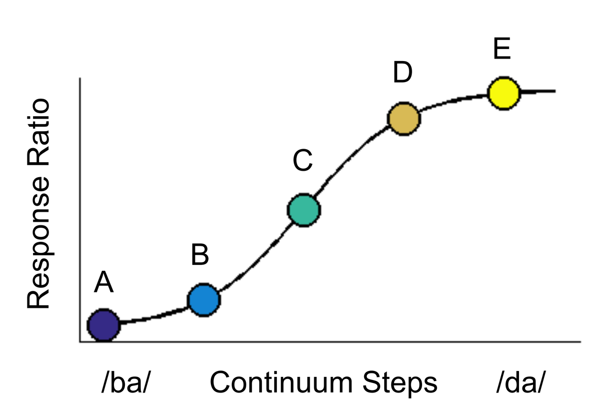
How does the brain respond to phonemic vs. phonetic auditory deviants?

Can phonemic vs. phonetic structure be decoded from patterns of neural activity?

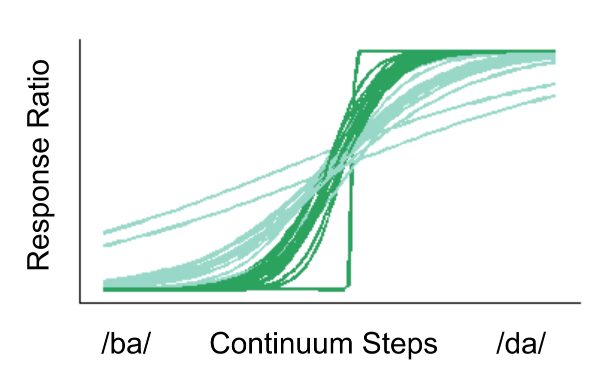
Participants were 33 adults (17 female) with normal hearing, defined as pure tone thresholds <25 dB at 0.25, 0.5, 1, 2, 4, 8 kHz. Thirteen met research criteria for dyslexia.

	Mean	SD	Range
Age (years)	26	6	19-41
Nonverbal IQ (KBIT)	111	14	86-132
Untimed word reading (WRMT)	101	12	77-118
Untimed pseudoword reading (WRMT)	92	16	59-121
Timed word reading (TOWRE)	100	14	73-130
Timed pseudoword reading (TOWRE)	95	14	61-115

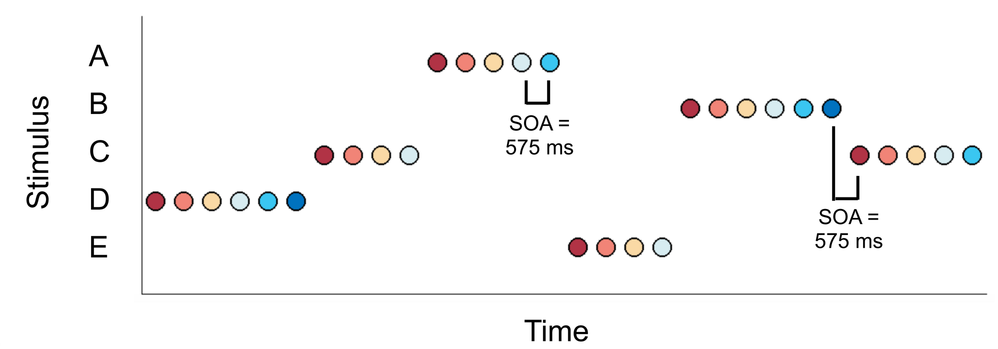
Stimulus Selection for the MEG paradigm occurred during another MEG task, in which the participant labeled tokens from a 10-step /ba/-/da/ acoustic continuum [2]. Forty trials of each step were administered in random order and the participant pressed a button to respond /ba/ or /da/. The task was self-paced with no feedback. We fit a logistic function to each individual's response ratio data. The step number nearest the inflection point was chosen as stimulus C, with A and B representing within-category /ba/ tokens and D and E within-category /da/ tokens. A, B, C, D, and E were equidistant.



Individual Differences in slopes were evident, irrespective of dyslexia/control status. We therefore performed some analyses on the whole group, and others on a median split (n=15 steep-slope participants and n=15 shallow-slope).



The **MEG Paradigm** was a roving-oddball design composed of "trains" of 4 to 6 repetitions of the same stimulus, A, B, C, D, or E. The first token of each train serves as a deviant, but, over time, becomes a standard for the next train. Transitions were balanced and pseudorandomized. Stimuli were delivered diotically via insert earphones. To maintain arousal, participants watched a soundless movie (Wall-E). A total of 3000 stimuli were presented in 28 minutes.

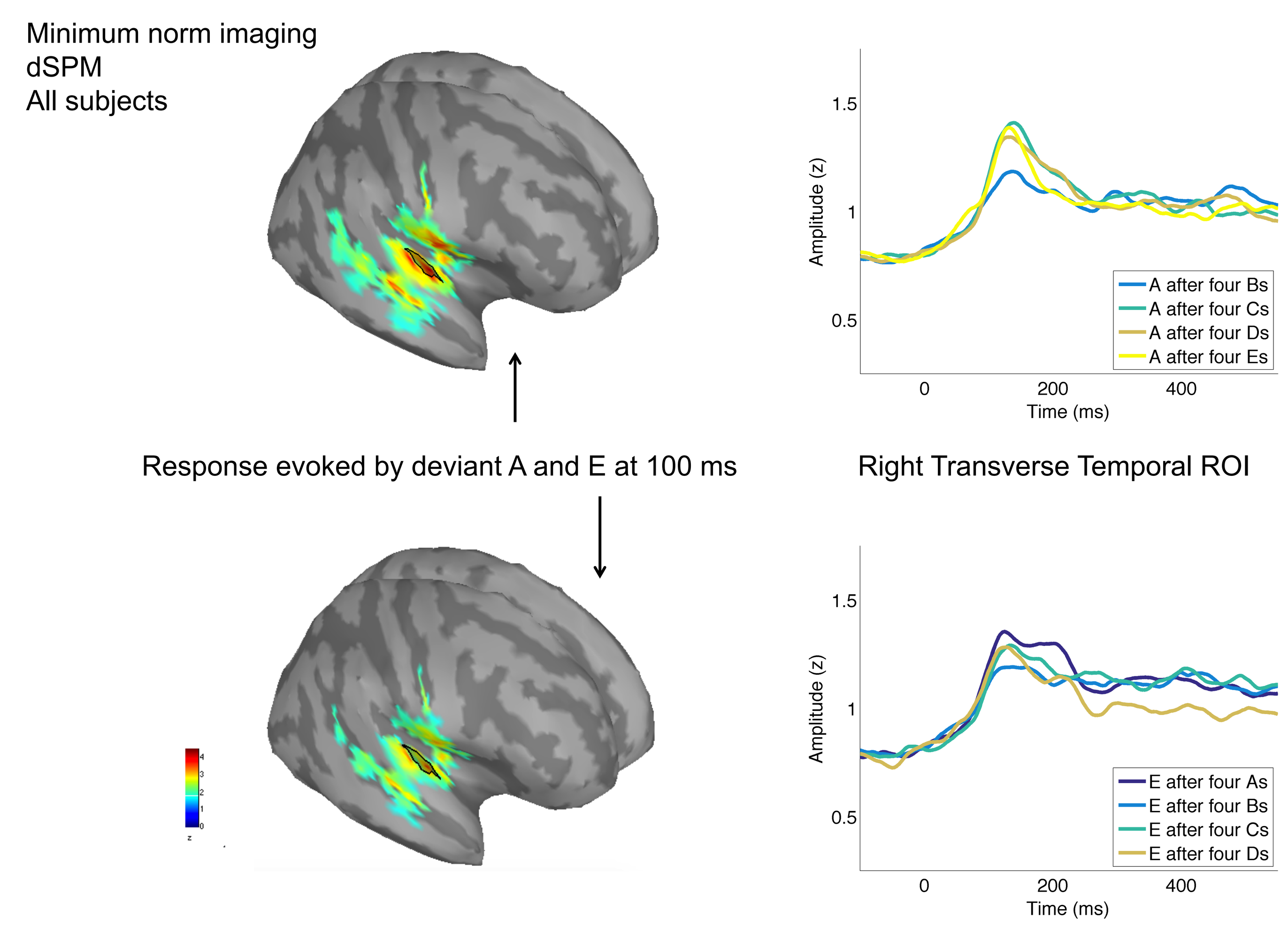


Recording of 306 MEG sensors was performed on an Elekta Neuromag TRIUX, with a sampling rate of 1 kHz and online filtering between 0.00003 and 0.33 kHz.
Preprocessing included using Maxfilter software to reduce noise and compensate for head movement. Brainstorm software was used to remove ocular artifacts via signal space projection, extract event-related epochs of -200 to 550 ms with respect to sound onset, and baseline-correct.
Source Estimation was performed with minimum norm imaging and dipole modeling approaches.

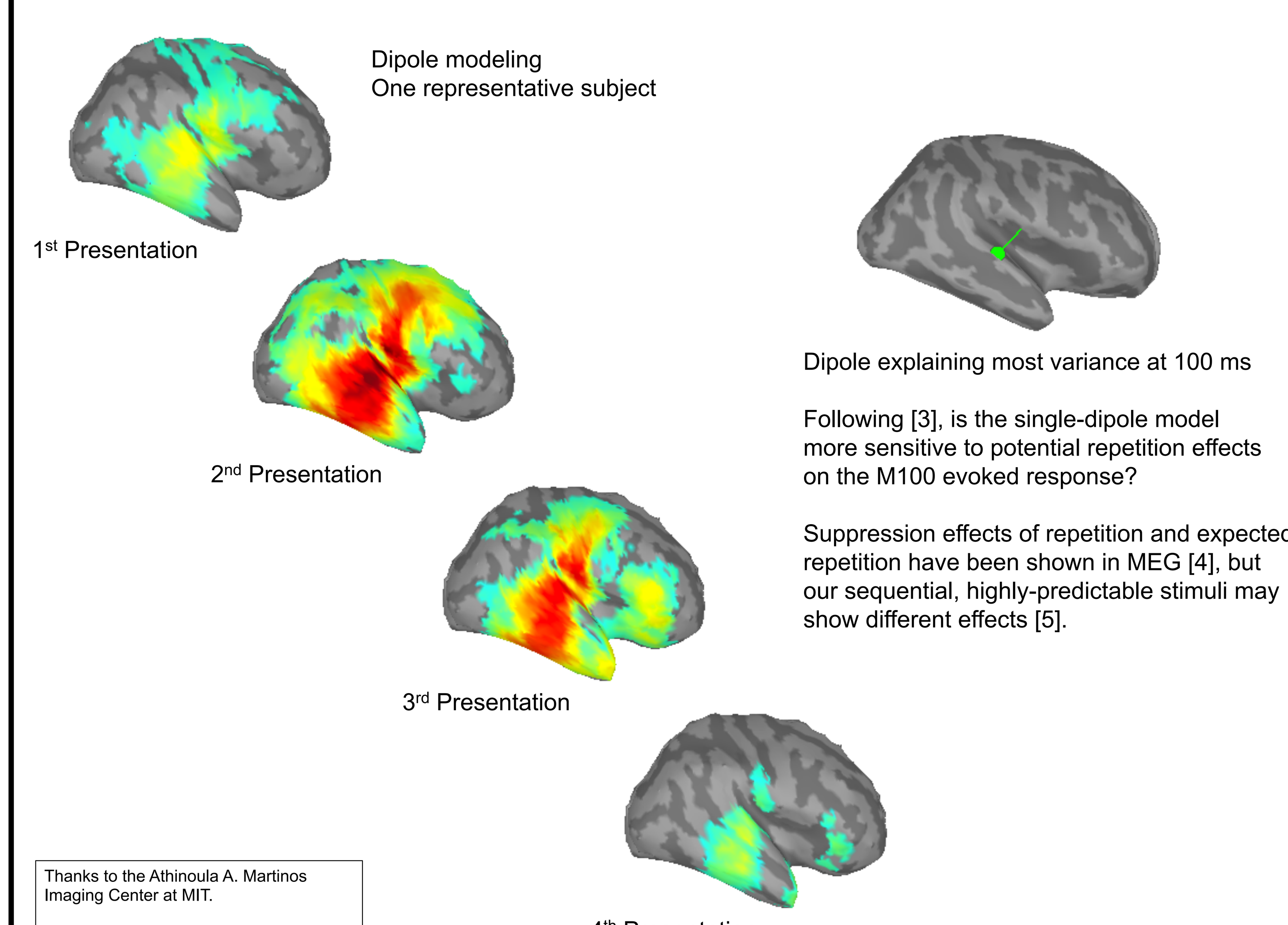
Multivariate Analysis involved training a multiclass support vector machine (SVM) to distinguish stimuli A, B, C, D, and E using LIBSVM software and custom MATLAB scripts. We performed decoding on low-pass-filtered (0.03 kHz) individual trial data in MEG sensor space at every timepoint. Channel baseline means were removed and standard deviations were set to 1. We implemented four-fold cross-validation with 100 permutations. Classification results were averaged over participants.

MEG Results

Effect of Preceding Standards on Deviant Response

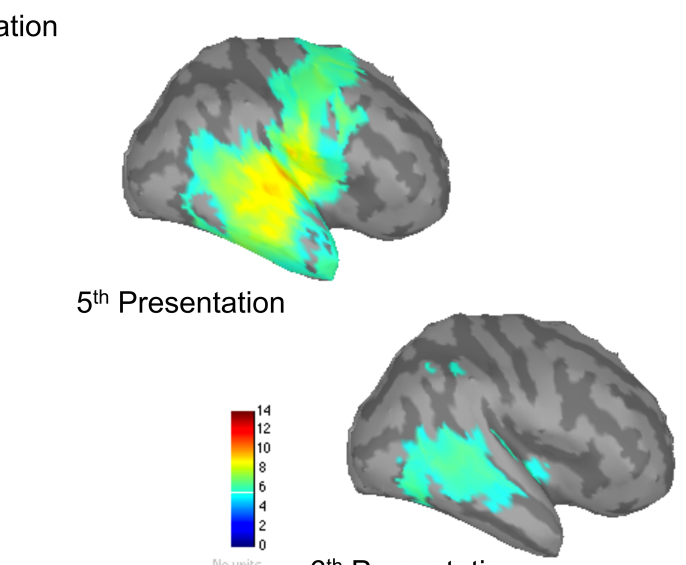


Effect of Repetition



Following [3], is the single-dipole model more sensitive to potential repetition effects on the M100 evoked response?

Suppression effects of repetition and expected repetition have been shown in MEG [4], but our sequential, highly-predictable stimuli may show different effects [5].



Thanks to the Athinoula A. Martinos Imaging Center at MIT.

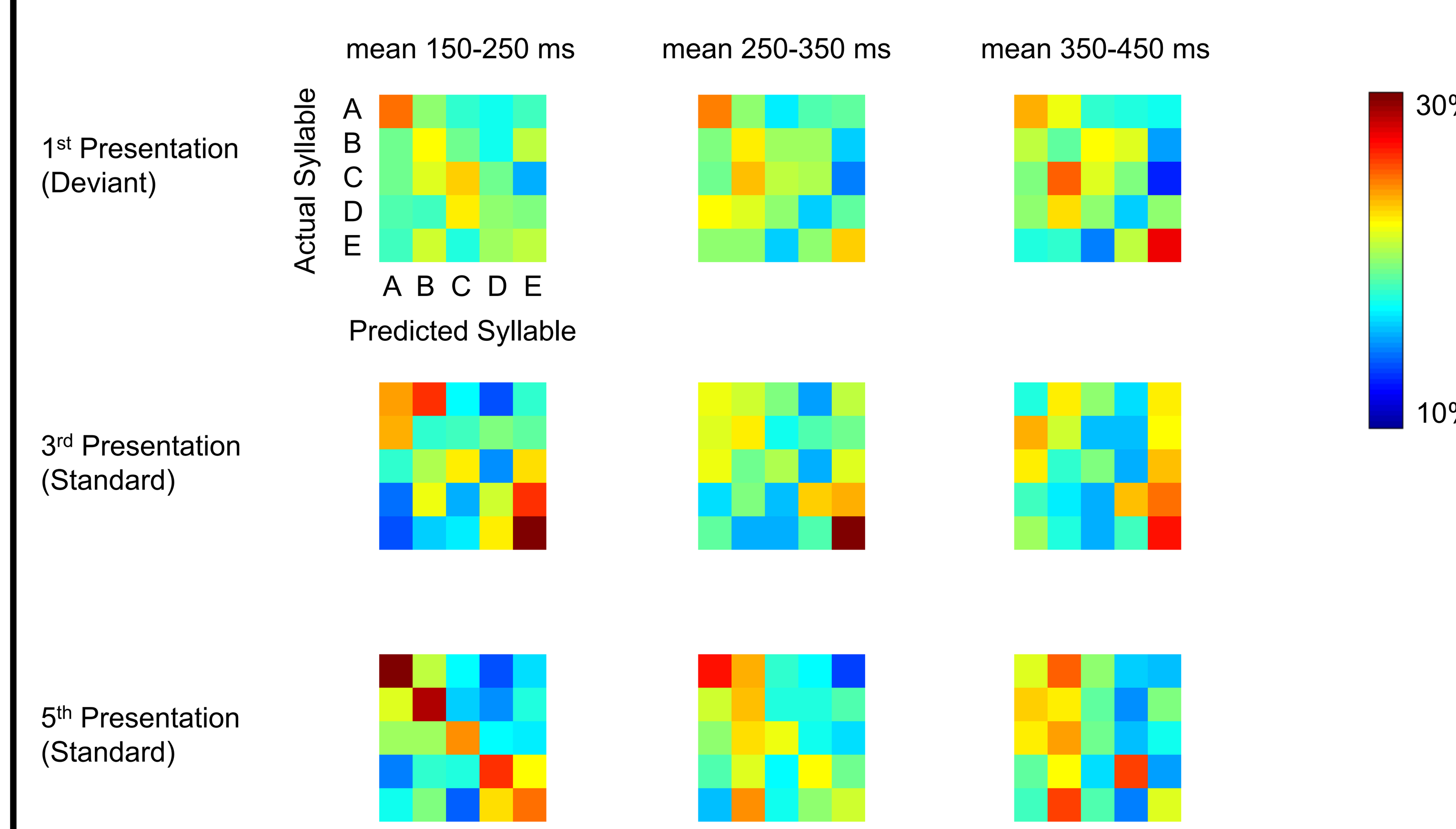
SDB is supported by a Friends of the McGovern Institute Graduate Fellowship.

contact: sdbeach@mit.edu

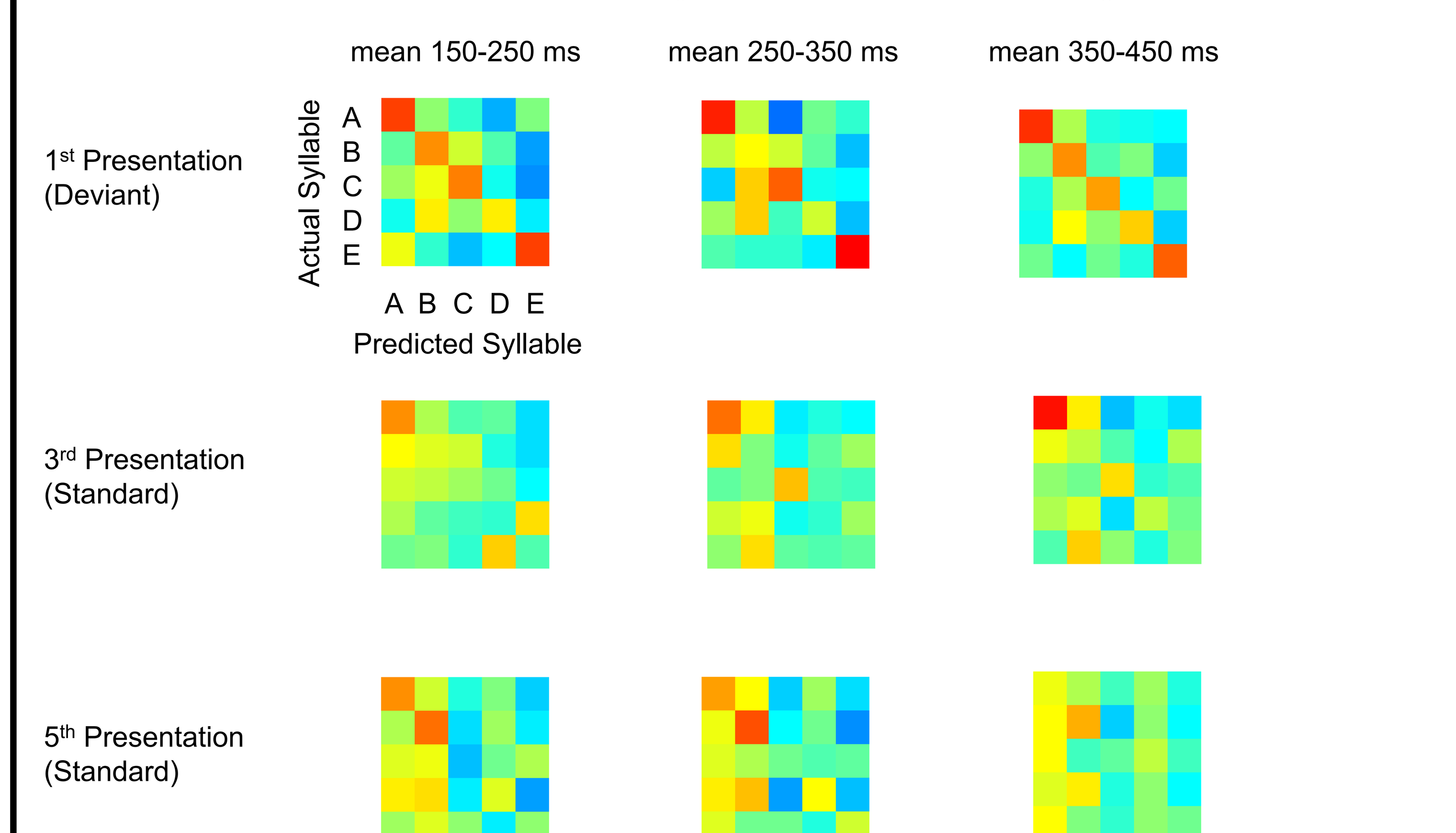
References
[1] Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Golestani, N., Wouters, J., & Ghesquière, P. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *PNAS, 107*(23), 10389-10394.
[2] Stephens, J. D., & Holt, L. L. (2011). A standard set of American-English voiced stop-consonant stimuli from morphed natural speech. *Speech communication, 53*(6), 877-888.
[3] Altmann, C.F., Uesaki, M., Ono, K., Matsuhashi, M., Mima, T., & Fukuyama, H. (2014). Categorical speech perception during active discrimination of consonants and vowels. *Neuropsychologia, 64*, 13-23.
[4] Todorovic, A., & de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *Journal of Neuroscience, 32*(39), 13389-13395.
[5] Hsu, Y. F., Hämäläinen, J. A., & Waszak, F. (2016). The auditory N1 suppression rebounds as prediction persists over time. *Neuropsychologia, 84*, 198-204.

Neural Decoding Results

Steep-Slope Participants: Emergence of Acoustic Decoding with Repetition



Shallow-Slope Participants: Decay of Acoustic Decoding with Repetition



Difference: Shallow - Steep