Visual cortical tracking of categorical speech features is enhanced by lip-reading training

Zhewei Cao[1], Aisling E. O'Sullivan[2], Lauren A. Szymula[1], Aaron R. Nidiffer[1], and Edmund C. Lalor[1,2]

[1]Department of Biomedical Engineering and Department of Neuroscience, University of Rochester, Rochester, NY, USA
[2]School of Engineering, Trinity Centre for Biomedical Engineering and Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland

Introduction

Neuroimaging has shown that during silent lip-reading, neural activity over the visual cortex encodes heirarchical visual speech signals. Visual EEG responses can be better predicted by using categorical speech features (i.e., visemes) in addition to low-level features like frame-to-frame motion (O'Sullivan et al., 2017; Hauswald et al., 2018).

Q1: Can trained lip-reading induce stronger tracking of visual speech?

2. Observing visual speech in the absence of auditory speech activates primary auditory cortex (Calver et al. 1997; Pekkola et al., 2005; Bourguinon et al., 2020). However it remains unclear what the activity reflects.

Q2: Is auditory cortical activation reflecting an ability to synthesize speech from visual features?

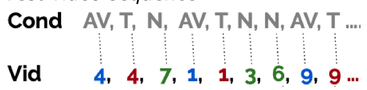
Methods

PARTICIPANTS
Sixteen native English speakers (11 females, ages 19-37) were recruited from the University of Rochester community.

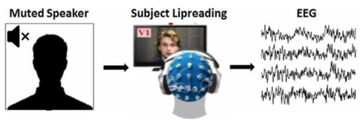
STIMULI & PROCEDURE
1. Training. Participants wer asked to watch five 1-minute long videos of a well-known speaker with intact sound. Videos were randomly selected from a set of 15 and presented 10 times each (50 presentations) in a randomized order

2. Testing. Participants were asked to perform a word detection task while watching muted versions of the five videos they were trained to lip-read, plus 5 more randomply selected from the set of 15. After each presentation, they were asked to subjectively rate its intelligibility.

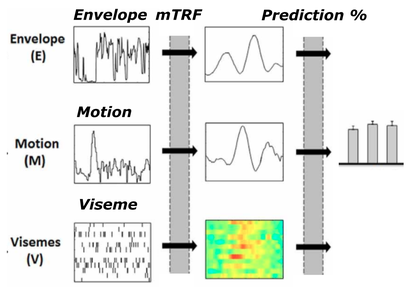
Test video sequence



EEG Data Acquisition



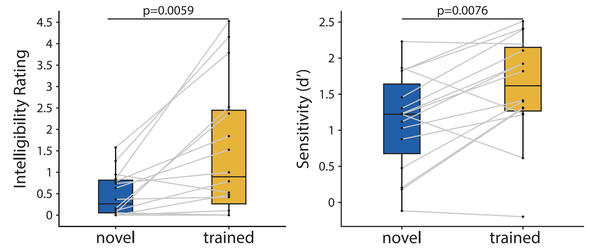
3. mTRF Analysis.



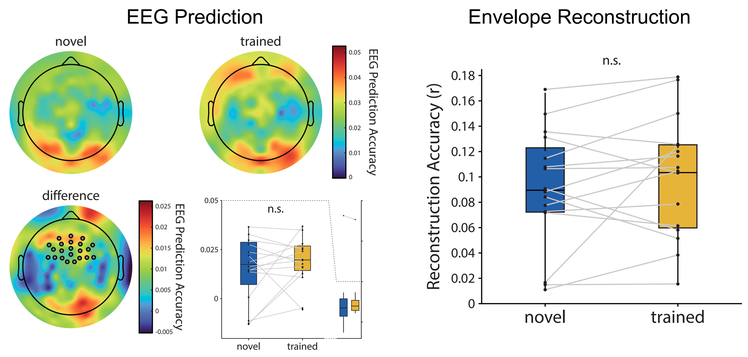
We mapped neural responses to the stimulus features using the multivariate Temporal Response Function Toolbox. With this analytical tool we could predict EEG responses from the stimuli (encoding model) or reconstruct the stimulus from EEG responses (decoding model).

Results

A. Training improved participants' lip-reading ability.



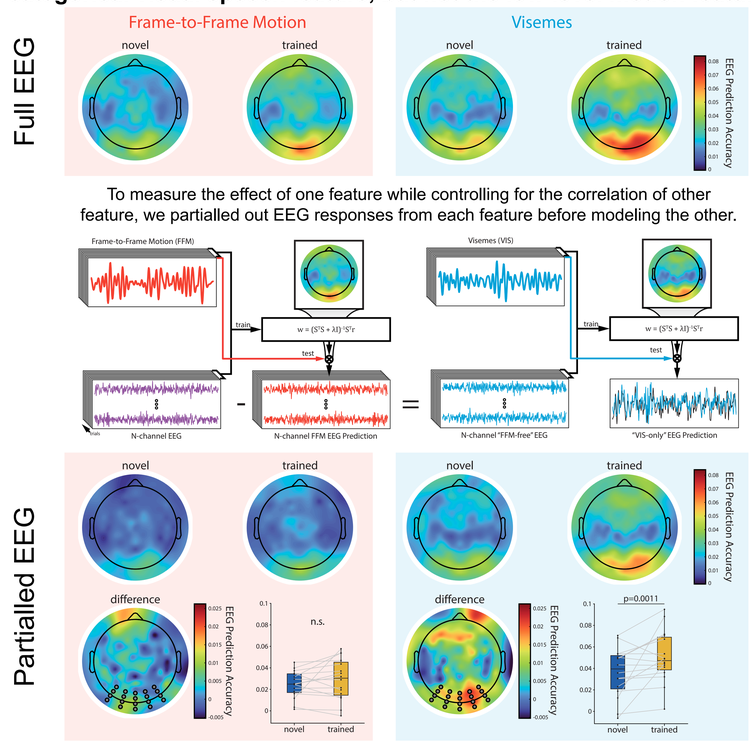
B. No improvement of the neural representation of the unheard speech envelope was observed after lip-reading improvements.



This was the case in both forward (EEG predition) modeling, when restricted to typical auditory (fronto-central) channels, and backward (envelope reconstruction) approaches.

Results

C. Lip-reading resulted in a more robust cortical encoding of the categorical visual speech feature, but not the low-level motion feature.



To measure the effect of one feature while controlling for the correlation of other feature, we partialled out EEG responses from each feature before modeling the other.

Summary

1. We found no supporting evidence for the enhanced encoding of the unheard acoustic envelope in typical auditory sensors.

2. Lip-reading improves visemic processing over occipital electrodes, suggesting an improvement in speech-specific - but not general - visual processing.

3. Future work: Isolate visual speech-specific features and examine their correlates to behavior (e.g., word detection).

References

Bernstein, L. E., & Liebenthal, E. (2014). Neural pathways for visual speech perception. Frontiers in Neuroscience, 8, 386.

Bourguignon, M., Baart, M., Kapnoula, E. C., & Molinaro, N. (2020) Journal of Neuroscience , 40(5), 1053–1065.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., Mcguire, P. K., et al. (1997). Science 276, 593–596.

Crosse, M. J., Butler, J. S., & Lalor, E. C. (2015). Journal of Neuroscience, 35(42), 14195-14204.

Crosse, M. J., Di Liberto, G. M., Bednar, A., and Lalor, E. C. (2016). Frontiers in Human Neuroscience 10:604.

Hauswald, A., Lithari, C., Collignon, O., Leonardelli, E., & Weisz, N. (2018) Current Biology: CB, 28(9), 1453-1459.e3

O'Sullivan, A. E., Crosse, M. J., Di Liberto, G. M., & Lalor, E. C. (2017). Frontiers in Human Neuroscience, 10, 679.

O'Sullivan, A. E., Lim, C. Y., & Lalor, E. C. (2019). European Journal of Neuroscience, 50(8), 3282–3295.

Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., Tarkiainen, A., et al. (2005). Neuroreport 16, 125–128.

Supported by NIH Grant R01 DC016297