Systems JC, OHSU - PREreview of "Cortical Representations of Speech in a Multi-talker Auditory Scene"

Daniela Saderi, Ph.D.[1]

[1]Oregon Health & Science University

May 5, 2020

Abstract

This is a preprint journal club review of **Cortical Representations of Speech in a Multi-talker Auditory Scene** by Krishna C Puvvada, Jonathan Z Simon. The preprint was originally posted on bioRxiv on April 10, 2017 (DOI: https://doi.org/10.1101/124750). The authors have responded to this review, and you can find the comments on bioRxiv. The article is now published in *The Journal of Neuroscience* (DOI: 10.1523/JNEUROSCI.0938-17.2017).
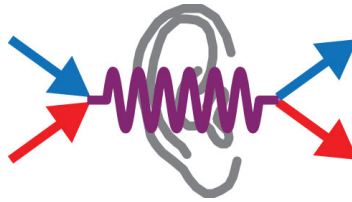
Figure 1: Systems Neuroscience JC - Oregon Health & Science University

Review

Dear authors,

Thank you for posting your work as a preprint on BioRxiv. We discussed your work at our latest (preprint) systems neuroscience journal club at the Oregon Health & Science University. Below is a summary of our feedback containing our main remarks, points of discussion, and suggestions.

This work explores the encoding and decoding properties of streaming auditory objects in primary and secondary regions of the human auditory cortex. Listeners were tasked to selectively attend to one of three overlapping speech streams while MEG activity was recorded. The novel aspect of the stimulus paradigm is the presence of three concomitant streams instead of two. Two main questions were explored: First, is the attended stream selectively represented in primary and/or secondary regions of the auditory cortex? Second, if that is the case, are the other two streams represented as a combined background, or are they represented separately as two distinct background streams?

To answer these questions two approaches were used: encoding and decoding strategies. The first suggestion we have about the work is purely organizational. We believe it would help the flow of the reading if the encoding and decoding analyses were described in the same order in the methods, the results (including figures), and the discussion sessions.

Regarding the encoding models, we think that it would be helpful to describe the degree of freedom of each model. We found it hard to evaluate if the newly proposed early-late model better describes the data when compared to the summation model because it truly captures key neuronal properties of auditory streaming, or if this is a result of the higher number of parameters. While cross-validation was employed during model fitting and addresses some of these concerns, a validation using an out-of-sample set for prediction would provide a definitive assessment of the potential existence and extent of overfitting.

While we appreciate that showing examples of the raw data together with the encoding and decoding model prediction contributes invaluably to clarity and transparency of the manuscript, we believe that figures 1 and 2 do not provide enough information for that to happen. Particularly, figure 2 would benefit from more labeling and perhaps even a more clear way of displaying the message.

In addition, it would be nice to see what the different models look like. What do the filter look like for the two encoding models? How do they compare? In line with the results of this study, one would believe that the late component of the late model should be larger, but without seeing the models it is hard to know for sure. Showing a representation of the model parameters would certainly add value to this work.

The following figures in the paper are definitely clearer and well accompany the text in the result session. However, we wondered, what does each point in the plots represent? Is it one data point per listener? Are they multiple data points per listener depending on the attended stimulus? It would really help if you could mention this at least in the caption. Perhaps adding a color coding would also help the reader better understand the results.

Finally, you mentioned that the 85-ms boundary was fit on a per-subject manner. Would it be possible to show a plot of those values to see how variable this boundary is in the sample data? We also wondered what it would happen to the model if you used the median of these values.

Thanks again for posting this work as a preprint. We really enjoyed discussing it at our JC and we hope these comments will help make the work even better.

Thank you,

Daniela Saderi (on behalf of the Systems JC, OHSU)