Dynamics of Pupil and Neural Responses to Contextual Speech Predictions

Sebastian Andric[1], Peter Donhauser[2], Sylvain Bailllet[3]

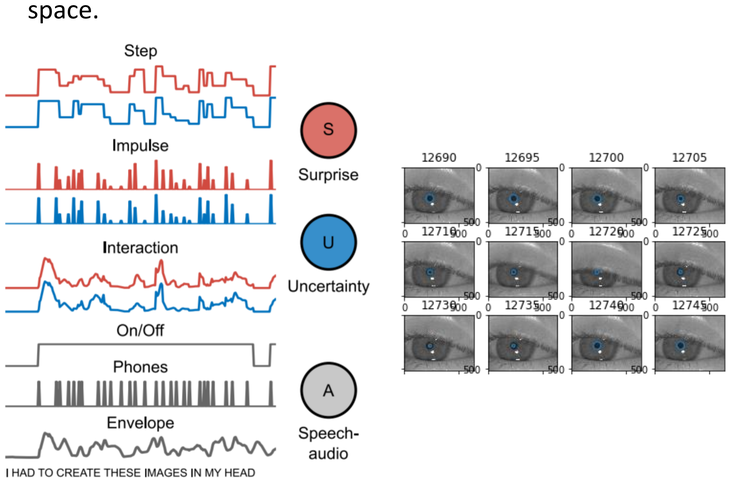
1. Integrated Program in Neuroscience, McGill;
2. Montreal Neurological Institute, McGill;
3. Department of Neurology and Neurosurgery, Biomedical Engineering and Computer Science, McGill

Introduction

- Incoming raw sensory content is diverse and dynamic, posing a challenging computational problem for brain processing[1].
- Internal representations of sensory contexts may alleviate this processing; these models modulate the gain of incoming information depending on whether they are expected or unexpected. Mechanistically, this is referred to as **predictive coding**[2].
- Human and animal studies show that **pupil diameter** is associated with attention and arousal[3], and has been shown to increase in size during presentation of **unpredictable audio events** and processing of complex auditory stimuli[4].
- The first goal of the project is to determine whether pupil dynamics are related to contextual speech predictions and to better understand the interaction between attentional and predictive processes.
- Some theories of predictive coding posit that the brain implements the predictive processing strategy through **neural oscillations** organized in different timescales[5] , where higher-order regions manifest slower dynamics.
- The second goal of the project is to determine whether auditory processing is organized within hierarchical neurophysiological activity during speech listening

Method

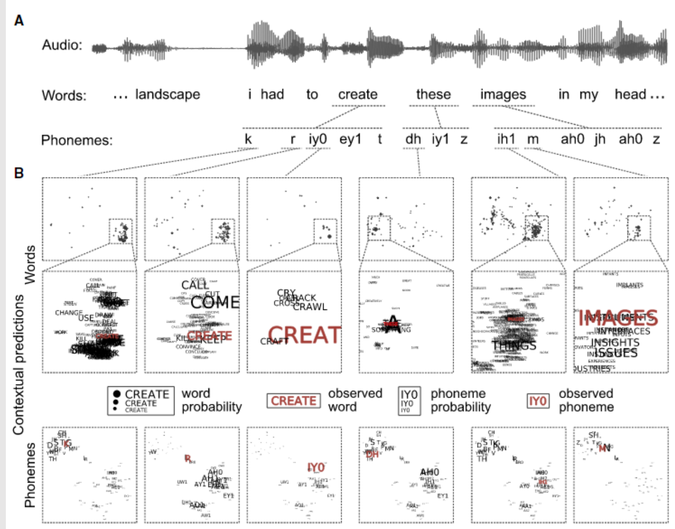
- Speech Dataset: The speech data used in neural network training and the magnetoencephalography (MEG) experiment were taken from the **TED-LIUM** corpus[6].
- Pupil Data: Pupil videos were taken during MEG recording, the Facemap toolbox was used for interpolation[7].
- Language Model: long short-term memory neural network, for complete specification see reference 1.
- Event Related Pupil Responses: Mean pupil response time=locked to highly surprising speech content.
- Ridge Regression and Cross Validation: Pupil dilation was predicted as a linear combination of the design matrix columns with penalty β. Leave-one-out cross validation was used to evaluate the model fit.
- Temporal Response Function (TRF): Regression weights averaged across regressors reshaped to produce a time-series sampled at temporal lags τ. TRFs were generated for a speech-audio feature space and the full model feature space.



From Donhauser and Baillet 2020

Results

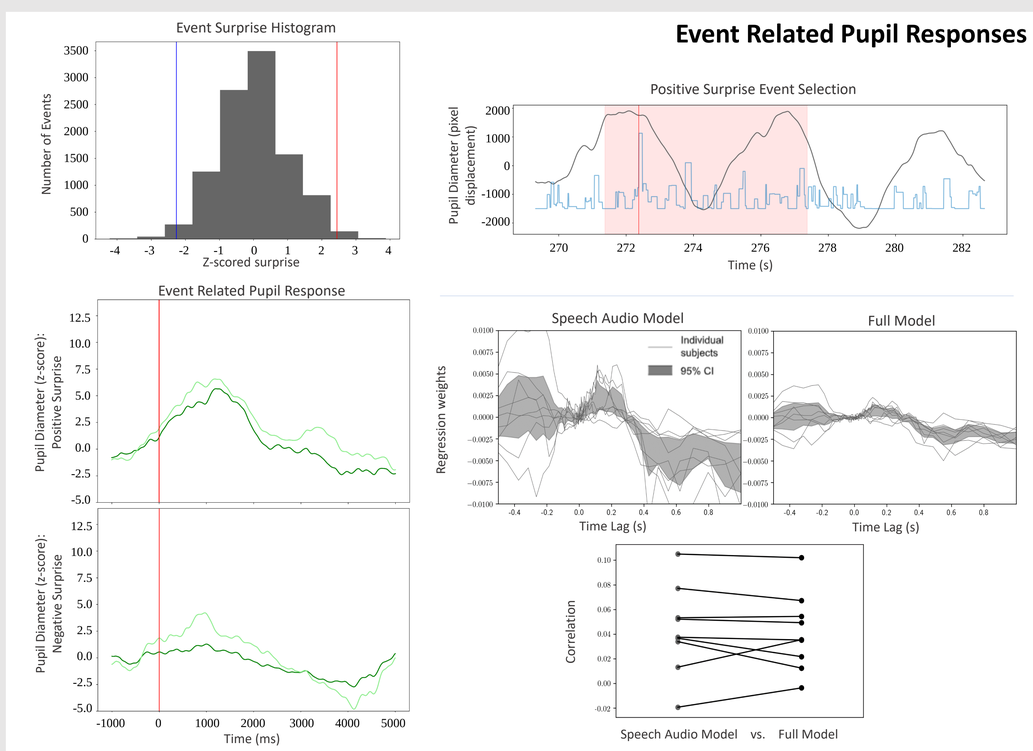
Contextual Speech Predictions



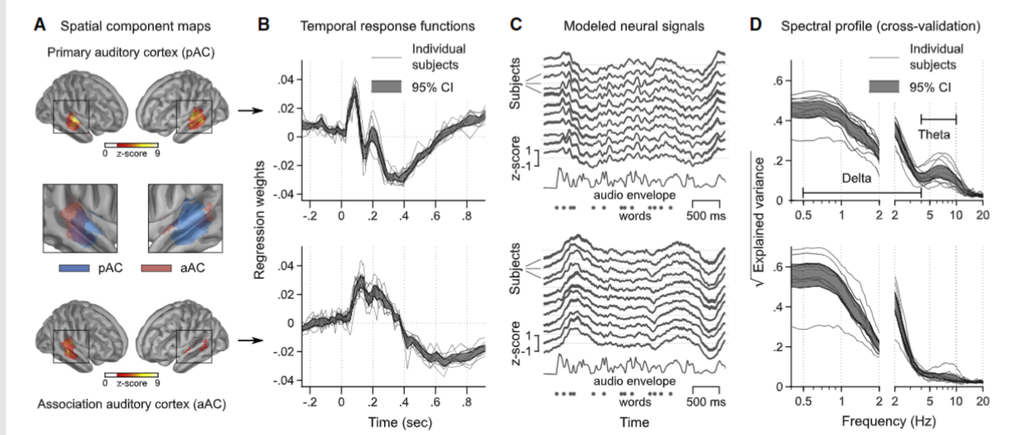
From Donhauser and Baillet 2020

- Contextual speech predictions generated by the neural network at the word- and phoneme-level.
- Contextual **uncertainty** is quantified by the entropy of the predictive distribution of the upcoming phoneme.
- **Surprise** is a measure of unexpectedness for the phoneme that is actually presented.
- We can analyze pupil responses and brain dynamics in response to the generated speech parameters

Event Related Pupil Responses



Speech-Related Neuophysiological Activity



From Donhauser and Baillet 2020

- A: Spatial components representing regions of activity well-explained by speech features
- B: TRFs averaged across all features (same as pupil TRF)
- C: Predicted neural responses from the regression model
- D: Regression model performance across the frequency spectrum

Discussion

- Contextual Speech Predictions: Predictions generated by the neural network have been previously shown to be associated with modulations in neural activity[1]. Here, we explored whether autonomic signals (pupil dilation) are also modulated by such predictions.
- Event Related Pupil Response: Pupil diameter increased during highly surprising speech segments, inline with previous literature describing pupil responses to unexpected events[4].
- Modeled Pupil Dynamics: The full contextual prediction model did not appear to improve predictions of the pupillary dynamics compared to the speech audio model. This might be due to noise in the traces of the pupil dynamics or overfitting introduced by the greater number of regressors.
- Modeled Neural Dynamics: The modeled neural responses reflect the differing timescales in neurophysiological activity; higher-order activity is better predicted at slower frequencies (delta) and downstream activity is better predicted at higher frequencies (theta)
- Current Directions: Different variations of the contextual predictions model will be assessed against the speech-audio model. Once a better suited model is identified, we may then investigate whether pupil dynamics are related to top down neural signals also related to predictions, such as beta oscillation bursting[8]. This may further illuminate the link between neural activations and autonomic functions.

References

1. Donhauser, P. W., & Baillet, S. (2020). Two Distinct Neural Timescales for Predictive Speech Processing. *Neuron*, 105(2). doi:10.1016/j.neuron.2019.10.019
2. Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79-87. doi:10.1038/4580
3. Mcginley, M., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C., . . . Mccormick, D. (2015). Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron, 87*(6), 1143-1161. doi:10.1016/j.neuron.2015.09.012
4. Zekveld, A. A., Koelewijn, T., & Kramer, S. E. (2018). The Pupil Dilation Response to Auditory Stimuli: Current State of Knowledge. *Trends in Hearing, 22*, 233121651877717. doi:10.1177/2331216518777174
5. Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. Neuron, 76(4), 695-711. doi:10.1016/j.neuron.2012.10.038
6. Rousseau, A., Deleglise, P., and Esteve, Y. (2014). Enhancing the TED-LIUM corpus with selected data for language modeling and more TED talks. In *LREC*, pp. 3935–3939
7. Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D. (2018). Spontaneous behaviors drive multidimensional, brain-wide activity. doi:10.1101/306019
8. Betti, V., Penna, S. D., Pasquale, F. D., & Corbetta, M. (2020). Spontaneous Beta Band Rhythms in the Predictive Coding of Natural Stimuli. *The Neuroscientist,* 107385842092898. doi:10.1177/1073858420928988

Acknowledgments

S.A. acknowledges the support from Caroline Palmer and the Complex Dynamics network through the National Science and Engineering Research Council of Canada – Collaborative Research and Training Experience Program grant. S.A. would also like to thank S.B., P,D. and the NeuroSPEED lab for their help and guidance.