Dissociating Predictive Sentence Processing from Acoustic Correlates in Naturalistic EEG Recordings from L1 and L2 Populations

The neural mechanism behind predictive processing is key to sentence comprehension. Prior work shows that the brain generates hierarchical predictions to gain information from upcoming inputs (Brennan & Hale, 2019; Heilbron et al., 2022; Schrimpf et al., 2021). Yet, linguistic predictability also robustly modulates the acoustic signal through which spoken language is expressed (e.g., Aylett & Turk, 2004) and the brain robustly responds to acoustic modulation (Brodbeck et al., 2022). It is thus unclear whether proposed neural signatures of predictive speech processing reflect unique cognitive processes or can be reduced to changes in acoustic energy. In this study, we aim to dissociate the effect of multiple levels of predictability from their acoustic correlates. To this end, we fit temporal response functions (TRFs; Brodbeck et al., 2023) against EEG data while participants listen to stories, using both acoustic information and linguistic annotations as predictors.

The TRF is a powerful tool to estimate neural response dynamics induced by both discrete and continuous features. In the current study, we fit TRFs against naturalistic EEG recordings with the following predictors: acoustic envelope, gammatone acoustic spectrogram, word frequency, next-word surprisal, and structural complexity from a bottom-up parser. We fit these models against two naturalistic EEG datasets: (1) the publicly available Alice in Wonderland EEG dataset (Bhattasali et al., 2020), and (2) a dataset recorded while English-L2 participants listened to the Little Prince audiobook in English; TRF analyses are identical between datasets. Gammatone predictors are generated from the original audiobooks and are believed to reflect the cochlear transformation of acoustic signals (Brodbeck et al., 2018). Envelopes are derived from the gammatone spectrograms by summing along the frequency dimension. Spoken word frequencies come from the SUBTLEX database. Surprisal for each word is computed as the negative log probability estimated by a GPT-2 model (Radford et al., 2019). Bottom-up parsing node counts follow that used in Brennan & Martin (2020).

We observe a significant modulation of acoustics by surprisal (Figure 1), such that higher surprisal values induce more acoustic energy in both the acoustic envelope and gammatone spectrogram. However, TRFs estimated with acoustic and linguistic predictors show that despite being intertwined, acoustic energy and word surprisal independently modulate the EEG signal: TRFs for each predictor explain unique variance (Figure 2) and they have different temporal dynamics. Surprisal TRFs show peaks in early rather than late time windows (see also Hale et al., 2018; cf. Frank et al., 2015); which may reflect the sensitivity of the methodologies used here. TRFs for word frequency explained additional variance and peak at around 350 ms, a time window commonly attributed to N400 (Lau et al., 2008). TRFs estimated for bottom-up node count point at a late frontal negativity peaking at around 500ms in line with prior accounts of late EEG components(Brennan & Martin, 2020; Hale et al., 2018; Kaan et al., 2000). Comparing the two datasets, the L2 TRFs for bottom-up node count have lower magnitude as well as a shorter time window, suggesting differences between L1 and L2 speakers.

In conclusion, we leverage TRFs to disentangle neural responses to acoustic energy and word predictability, which are highly intertwined, along with word frequency and node counts derived from a bottom-up parser. Our results suggest that word frequency, linearly derived surprisal, and bottom-up parsing show distinct temporal signatures alongside acoustic covariates.

Figure 1. Estimated acoustic modulations of GPT-2 surprisal.

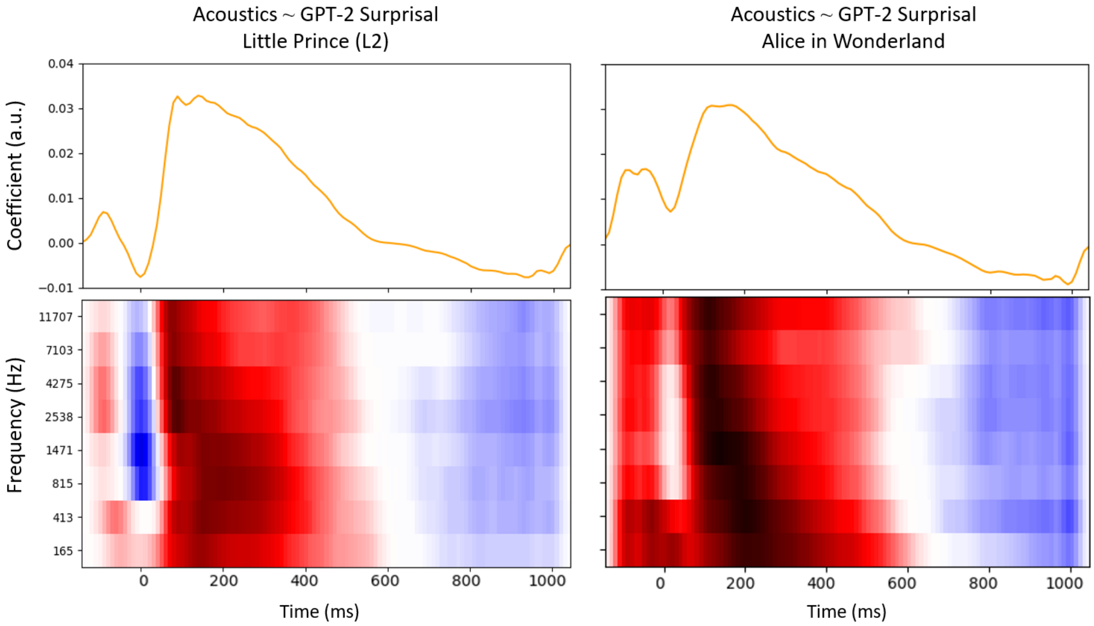
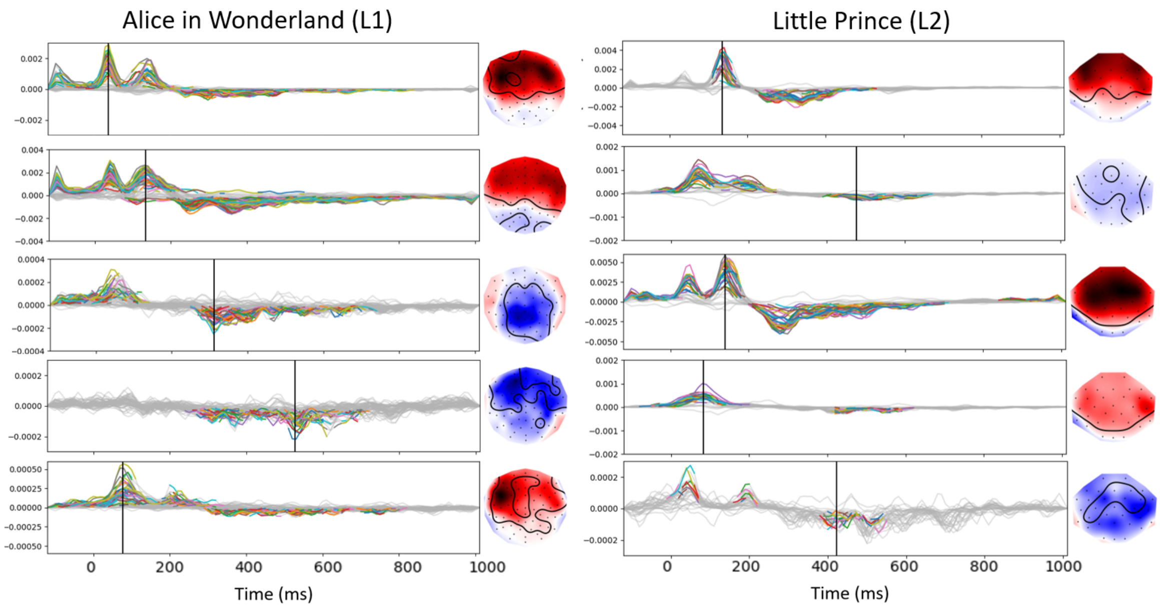


Figure 2. Estimated temporal response functions for all predictors on both EEG datasets.



References

Aylett, M., & Turk, A. (2004). The Smooth Signal Redundancy Hypothesis: A Functional Explanation for Relationships between Redundancy, Prosodic Prominence, and Duration in Spontaneous Speech. *Language and Speech*, *47*(1), 31–56. https://doi.org/10.1177/00238309040470010201

Bhattasali, S., Brennan, J., Luh, W.-M., Franzluebbers, B., & Hale, J. (2020). The Alice Datasets: fMRI & EEG Observations of Natural Language Comprehension. In N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 120–125). European Language Resources Association. https://aclanthology.org/2020.lrec-1.15

Brennan, J. R., & Hale, J. T. (2019). Hierarchical structure guides rapid linguistic predictions during naturalistic listening. *PLOS ONE*, *14*(1), e0207741. https://doi.org/10.1371/journal.pone.0207741

Brennan, J. R., & Martin, A. E. (2020). Phase synchronization varies systematically with linguistic structure composition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1791), 20190305. https://doi.org/10.1098/rstb.2019.0305

Brodbeck, C., Bhattasali, S., Cruz Heredia, A. A., Resnik, P., Simon, J. Z., & Lau, E. (2022). Parallel processing in speech perception with local and global representations of linguistic context. *eLife*, *11*, e72056. https://doi.org/10.7554/eLife.72056

Brodbeck, C., Das, P., Gillis, M., Kulasingham, J. P., Bhattasali, S., Gaston, P., Resnik, P., & Simon, J. Z. (2023). Eelbrain, a Python toolkit for time-continuous analysis with temporal response functions. *eLife*, *12*, e85012. https://doi.org/10.7554/eLife.85012

Brodbeck, C., Hong, L. E., & Simon, J. Z. (2018). Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech. *Current Biology*, *28*(24), 3976-3983.e5. https://doi.org/10.1016/j.cub.2018.10.042

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, *140*, 1–11. https://doi.org/10.1016/j.bandl.2014.10.006

Hale, J., Dyer, C., Kuncoro, A., & Brennan, J. R. (2018, June 11). *Finding Syntax in Human Encephalography with Beam Search*. arXiv.Org. https://arxiv.org/abs/1806.04127v1

Heilbron, M., Armeni, K., Schoffelen, J.-M., Hagoort, P., & de Lange, F. P. (2022). A hierarchy of linguistic predictions during natural language comprehension. *Proceedings of the National Academy of Sciences*, *119*(32), e2201968119. https://doi.org/10.1073/pnas.2201968119

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201. https://doi.org/10.1080/016909600386084

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), Article 12. https://doi.org/10.1038/nrn2532

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). *Language Models are Unsupervised Multitask Learners*.

https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe

Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2021). The neural architecture of language: Integrative modeling converges on predictive processing. *Proceedings of the National Academy of Sciences*, *118*(45), e2105646118. https://doi.org/10.1073/pnas.2105646118