Prediction in the language network is sensitive to syntactic structure

Cory Shain[1], Idan Blank[2], Marten van Schijndel[3], Evelina Fedorenko[4], William Schuler[1] [1]OSU, [2]UCLA, [3]Cornell, [4]MIT

According to many models of human sentence processing, the mind incrementally computes syntactic representations during language comprehension and uses them to guide expectations about upcoming words [13, 17, 24]. This position is supported by experimental evidence of responses to syntactic manipulations, both in the reading record [20, 12, 3] and in neural measures [19, 9, 23, 1, 21], including manipulations that directly target syntactic influences on expectations [30]. However, evidence from naturalistic studies for syntax-sensitivity in participants' expectations about future words has been more mixed: while some studies have found evidence that probabilities generated by a probabilistic context free grammar (PCFG) parser explain variance in the human sentence processing response [14, 25, 5, 32, 2, 16], others have argued that these effects can largely be explained on the basis of surface-level word co-occurrence statistics alone and have questioned the extent to which everyday sentence comprehension relies on hierarchical syntactic representations [6, 7, 8]. A recent large-scale naturalistic fMRI study showed that contextual predictability modulates the neural response in left-hemisphere language (LANG) regions but not in bilateral multiple demand (MD) regions [27]. However, their 5-gram predictability estimate emphasizes lexical co-occurrence patterns at the expense of syntactic structure, since it cannot directly model grammatical categories, syntactic composition, or dependencies longer than 5 words. In this study, we reanalyze the data from [27] and show that PCFG estimates of word predictability explain substantial variance in the neural response above 5-grams alone, suggesting that participants did exploit top-down estimates of syntactic structure to guide interpretations and supporting the hypothesis that such structures are computed rapidly and online during everyday sentence processing. As in [27], these effects are restricted to the language network, not the multiple demand network, suggesting that even top-down structural constraints are locally encoded in the language circuits, rather than relying on domain-general executive control.

The dataset analyzed in [27] contains fMRI timecourses from 78 participants (30 males) who first completed a localizer task designed to identify participant-specific functional regions of interest (fROIs): six left-hemisphere fronto-temporal language areas and twenty bilateral fronto-parietal multiple demand areas. Participants then listened to audio presentation of materials from the Natural Stories corpus [10]. Neural measures during story listening were recorded by an fMRI scanner and analyzed using continuous-time deconvolutional regression (CDR) [26, 29], a technique that estimates continuous hemodynamic response functions (HRFs) directly from naturalistic data. CDR models include a fixed and by-fROI slope for *TR number* (scan index), HRF shape parameters by fROI, and fixed and by-fROI HRF amplitudes for both (1) control variables *rate* (deconvolutional intercept term), *sound power* (root mean squared energy of audio stimuli), and *word frequency*, as well as (2) critical variables *5-gram surprisal* and *PCFG surprisal*. Word frequency and 5-gram models are estimated using KenLM [15] trained on Gigaword 3 [11], while PCFG models are estimated using an incremental left-corner parser [31] trained on sections 2 through 21 of the Wall Street Journal corpus [18] reannotated into a cognitively-inspired generalized categorical grammar [22]. The explanatory contribution of both surprisal measures is assessed using an ablative bootstrap test of likelihood in a held-out dataset, quantifying the degradation in generalization performance from removing fixed effects for one or both surprisal measures.

As shown in Figure 1 and Table 1, both *5-gram surprisal* and *PCFG surprisal* significantly and independently modulate the neural response in the language network, while neither significantly modulates the neural response in the multiple-demand network. A further ablative comparison of combined models of responses in both regions shows that the difference between the networks is significant. In addition, by-fROI analyis shows that the LANG model explains held-out variance in 5/6 fROIs, while the MD model explains held-out variance in only 1/20 fROIs (see [28] for more detailed analysis of individual fROIs). Together, these results align with studies cited above in supporting the existence of predictive mechanisms for language processing that compute and condition on syntactic structures during ordinary sentence processing, not just under constructed syntactic manipulations. Furthermore, they show that these mechanisms reside primarily in neural circuits that specialize for language processing, rather than in domain-general multiple-demand circuits whose activity scales with overall cognitive load [4]. This finding illuminates prior arguments that language processing relies on a domain-general and even species-general capacity for hierarchic sequential prediction [31, 24]: our study suggests that, at least in the case of hierarchical language structure, this general capacity is nevertheless locally implemented in specialized neural language circuits.

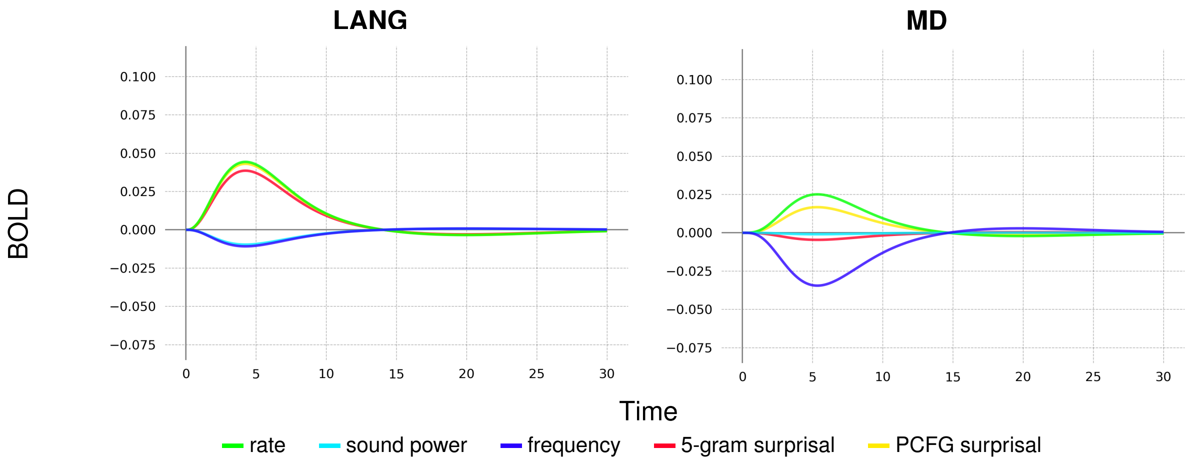

Figure 1: Estimated hemodynamic response functions by network. Although the *PCFG surprisal* estimate is positive in MD, it is not significant (Table 1) and the full model explains no held-out variance.

	LANG		MD	
Comparison	Estimate	p	Estimate	p
5-gram over neither	0.307	**0.0001*****	0.019	0.137
PCFG over neither	0.352	**0.0001*****	0.081	1.0
5-gram over PCFG	0.209	**0.0001*****	-0.025	1.0
PCFG over 5-gram	0.235	**0.0001*****	0.097	1.0

Table 1: Held-out permutation testing results. *Estimate* comes from the full model in each comparison.

References

[1] Brennan, J. and Pylkkänen, L. *NeuroImage*, 2012.
[2] Brennan, J., Stabler, E. P., Van Wagenen, S. E., Luh, W.-M., and Hale, J. T. *Brain and Language*, 2016.
[3] Clifton Jr, C. and Staub, A. *Eye*, 2011.
[4] Duncan, J. *Trends in Cognitive Sciences*, 2010.
[5] Fossum, V. and Levy, R. In *Proceedings of the 3rd Workshop on Cognitive Modeling and Computational Linguistics*, 2012.
[6] Frank, S. L. and Bod, R. *Psychological Science*, 6 2011.
[7] Frank, S. L., Bod, R., and Christiansen, M. H. *Proceedings of the Royal Society B: Biological Sciences*, 2012.
[8] Frank, S. L. and Christiansen, M. H. *Language, Cognition and Neuroscience*, 2018.
[9] Friederici, A. D., Rueschemeyer, S.-A., Hahne, A., and Fiebach, C. J. *Cerebral Cortex*, 2003.
[10] Futrell, R., Gibson, E., Tily, H. J., Blank, I., Vishnevetsky, A., Piantadosi, S., and Fedorenko, E. In *LREC*, 2018.
[11] Graff, D., Kong, J., Chen, K., and Maeda, K. English Gigaword Third Edition LDC2007T07, 2007.
[12] Grodner, D. J. and Gibson, E. *Cognitive Science*, 2005.
[13] Hale, J. In *NAACL*, 2001.
[14] Hale, J. *Cognitive Science*, 7 2006.
[15] Heafield, K., Pouzyrevsky, I., Clark, J. H., and Koehn, P. In *ACL*, 2013.
[16] Henderson, J. M., Choi, W., Lowder, M. W., and Ferreira, F. *Neuroimage*, 2016.
[17] Levy, R. *Cognition*, 2008.
[18] Marcus, M. P., Santorini, B., and Marcinkiewicz, M. A. *Computational Linguistics*, 1993.
[19] Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., and Mehler, J. *Journal of Cognitive Neuroscience*, 1993.
[20] Mehler, J., Bever, T. G., and Carey, P. *Perception & Psychophysics*, 1967.
[21] Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., Cash, S. S., Naccache, L., Hale, J. T., Pallier, C., and others. *Proceedings of the National Academy of Sciences*, 2017.
[22] Nguyen, L., van Schijndel, M., and Schuler, W. In *Proceedings of COLING 2012*, 2012.
[23] Pallier, C., Devauchelle, A.-D., and Dehaene, S. *Proceedings of the National Academy of Sciences*, 2011.
[24] Rasmussen, N. E. and Schuler, W. *Cognitive Science*, 2018.
[25] Roark, B., Bachrach, A., Cardenas, C., and Pallier, C. In *EMNLP*, 2009.
[26] Shain, C. and Schuler, W. In *EMNLP*, 2018.
[27] Shain, C., Blank, I., van Schijndel, M., Schuler, W., and Fedorenko, E. In *CUNY 2019*, 2019.
[28] Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., and Fedorenko, E. *Neuropsychologia*, 2019.
[29] Shain, C. and Schuler, W. *PsyArXiv*, 2019.
[30] Staub, A. and Clifton, C. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2006.
[31] van Schijndel, M., Exley, A., and Schuler, W. *Topics in Cognitive Science*, 2013.
[32] van Schijndel, M. and Schuler, W. In *Proceedings of NAACL-HLT 2015*, 2015.