Tracking the emergence of meaning in the brain during natural story comprehension

Phillip Alday   Jona Sassenhagen   Ina Bornkessel-Schlesewsky

Philipps Universität Marburg

University of South Australia

Cognitive Neuroscience Laboratory

Introduction

Brain activity indexes a word's contextual fit into a rich model based on prior experiences [1, 2] and has been characterised as reflecting the spreading activation of semantic features in long-term memory [3, 4]. Following [5], we created a pseudo Bayesian model of contextual support using data from a large corpus [6]. Context is infinite, yet is exponentially dampened [7], thus near or oft repeated contextual support dominates the activation function.

Experiment

52 subjects listened to a short story. EEG was cleaned of artifacts with ICA [8] and 52×1682 segments extracted, time locked to the onset of content words.

Primitives

Semantic association can be modelled with a few primitives based on statistical regularities of the input.

$\mathrm{cooc}(x, y)$ = Shared co-occurrence of two words

$\mathrm{sig}(x, y)$ = Significance of shared co-occurrence of two words (measure of co-occurrence frequency)

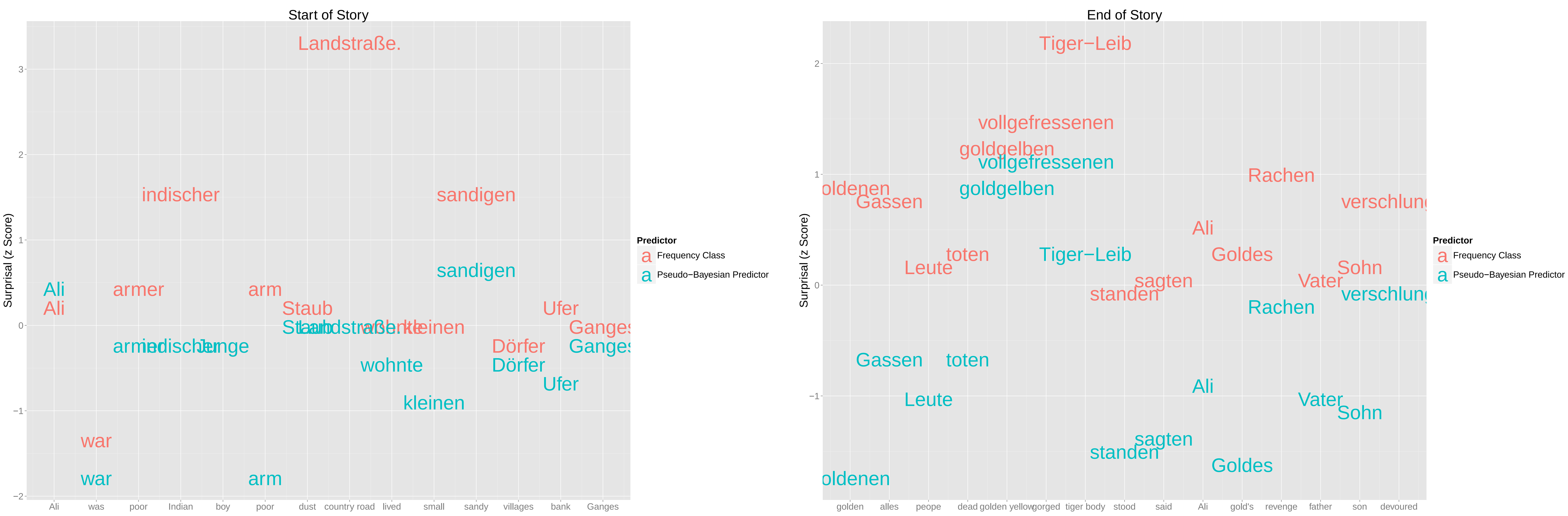
$\omega(x)$ = Logarithmic frequency class (bigger is rarer=more surprising)

$\mathcal{I}(x)$ = index of most recent occurrence of lexeme

$\mathcal{N}(x)$ = number of previous occurrences of lexeme

$\mathcal{NC}(x)$ = number of previous lexemes for which this lexeme is a co-occurrence

Predictor vs. Logarithmic Frequency Class



Frequency and the pseudo-Bayesian predictor are initially similar, reflecting the dominance of baseline frequency in the absence of a rich context. By the end of the story, however, their divergence becomes more pronounced as the amount of prior information available has greatly increased.

Contextual Support (Semantic Priming)

surprisal based on corpus frequency and semantic-contextual compatibility

Exponential dampening

scaling

Frequency interaction (geometric mean)

$$P(x_n, m) = \omega(x_n) - \log\left(\sum_{k=1}^{m} e^{-\sqrt{k}} \frac{\sqrt{\omega(x_n)\omega(x_{n-i})}}{|\mathrm{cooc}(x_n, x_{n-i})|}\right)$$

Length of context window

Number of shared co-occurrences (reduction in surprise)

Repetition Priming

Repetition priming reduces amplitude

Decay of effect with time since last seen

$$R(x_n) = 1 + \lfloor \log(n - \mathcal{I}(x_n) + 1) \rfloor - \lceil \log_2 \mathcal{N}(x_n) \rceil$$

$$P'(x_n, m) = \min\{R(x_n), P(x_n, m)\}$$

P prime(d)

Additional repetition strengthens the effect

Decay can't increase amplitude beyond baseline

Pseudo-Bayesian Prediction (Lexeme-specific prediction)

Prior probability of a specific lexeme further reduces amplitude

Baseline co-occurence frequency increases expectation

$$\Pi(x_n) = \log_{10}\left(1 + \sum_{\substack{y \in \mathrm{cooc}(x_n, x_i) \\ i < n}} \mathrm{sig}(x_n, y)\mathcal{NC}(y)\right)$$

$$P''(x_n, m) = \lceil P'(x_n, m) - \Pi(x_n) \rceil$$

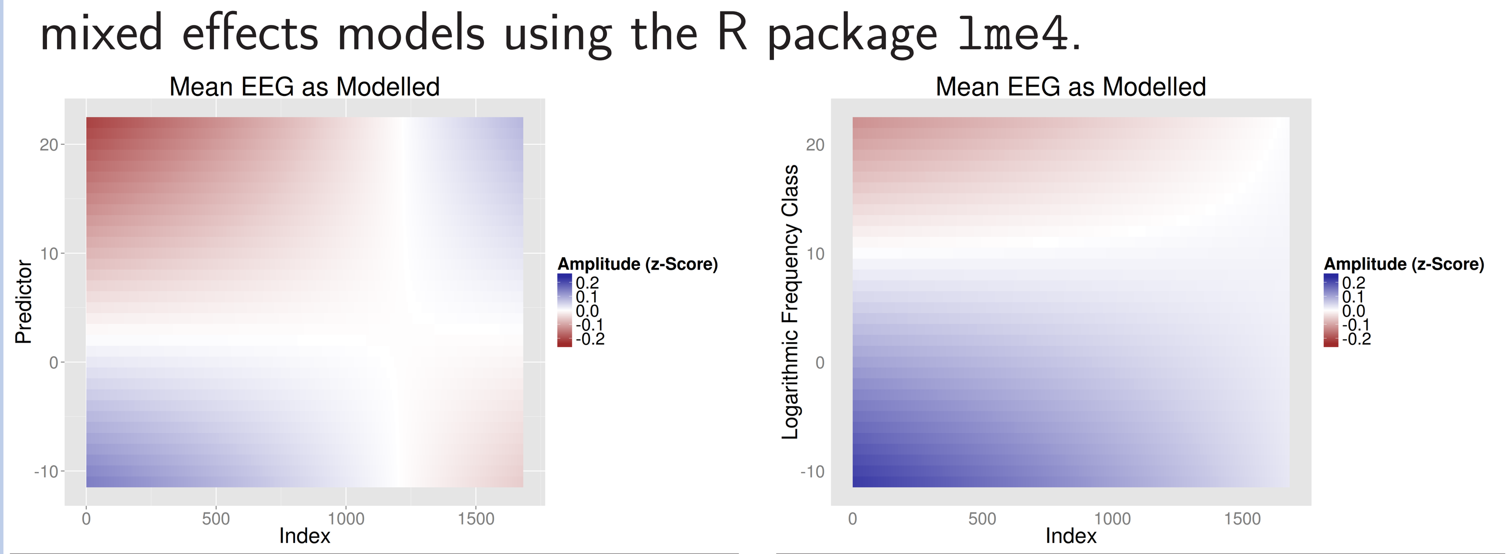
Repeated co-occurence increases expectation

Examine all previous co-occurences

P double prime (with 2nd-order concurrence prior)

Models

Predictors were compared to scaled, single-trial mean EEG via mixed effects models using the R package `lme4`.



	Estimate	Std. Error	t value
(Intercept)	0.012	0.0088	1.4
index	-2.2e-05	8.5e-06	-2.6
pred	-0.014	0.0011	-13
index:pred	1.3e-05	1.1e-06	12

	Estimate	Std. Error	t value
(Intercept)	0.11	0.017	6.4
index	-9.4e-05	1.8e-05	-5.2
freq	-0.013	0.0012	-11
index:freq	1e-05	1.3e-06	8

(ROI and its interactions omitted)

Semantic support correlated negatively with EEG amplitude in the N400 time window. While frequency is a decent global prior, the stronger interaction of our predictor with index allows for better modelling of the constrained local context.

Conclusion

Explicit computational models allow the prediction of neural activity corresponding to semantic processing for naturalistic stimuli, beyond categorical designs and highly artificial task contests. Our pseudo-Bayesian model of semantic expectability interacts dynamically with context during story comprehension. The complex semantics of coherent narratives result in more complex neural patterns than observed in classical experiments.

Literature

[1] K. J. Friston (2005). *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*.
[2] M. S. George, S. Mannes, et al. (1994). *Journal of Cognitive Neuroscience*.
[3] S. Laszlo & D. C. Plaut (2012). *Brain and language*.
[4] J. L. Elman (2004). *Trends in Cognitive Science*.
[5] M. M. Kutas & K. D. Federmeier (2011). *Annual review of psychology*.
[6] U. Quasthoff, M. Richter, et al. (2006).
[7] J. Jonides, R. L. Lewis, et al. (2008). *Annual review of psychology*.
[8] I. Winkler, S. Haufe, et al. (2011). *Behavioral and Brain Functions*.
[9] D. Bates, M. Maechler, et al. (2014). *lme4: Linear mixed-effects models using S4 classes*. R package version 1.1-7.