Neural signatures of incremental text processing correlate with word entropy in a natural story context

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 characterized as reflecting the spreading activation of semantic features in long-term memory [3, 4]. Here, we investigate the emergence of meaning in a non-categorical, natural, coherent design.

Information Theory

Information theory was originally presented as a mathematical theory of communication [5] and one of its first applications was predictive codes for language [6]. Yet its use in modern linguistics is limited to statistical applications for model comparison [7, 8], an oracle in computational linguistics [9, 10] and some attempts at an automatic diagnostic tool in neuropsychiatry [11]. Although some attempts have been made at integrating quantitative measures of information content with psychology and linguistic theory [12, 13], no study thus far has combined such measures with neurophysiological data collected during naturalistic language processing.

Entropy, Information and the N400

Researchers of language have identified brain signatures (e.g. the N400) of the fit of a word with its context, with more context providing a better fit [14, 15]. Recent models have postulated that the N400 reflects integration costs of top-down expectations with bottom-up stimulus features [16] and in general indexes fulfilment of expectation. Entropy measures the average new information content or "disorder" introduced by a particular symbol.

probability of a particular word (relative frequency)

information content

$$H = -\frac{1}{\log n} \sum_{w} p(w) \log p(w)$$

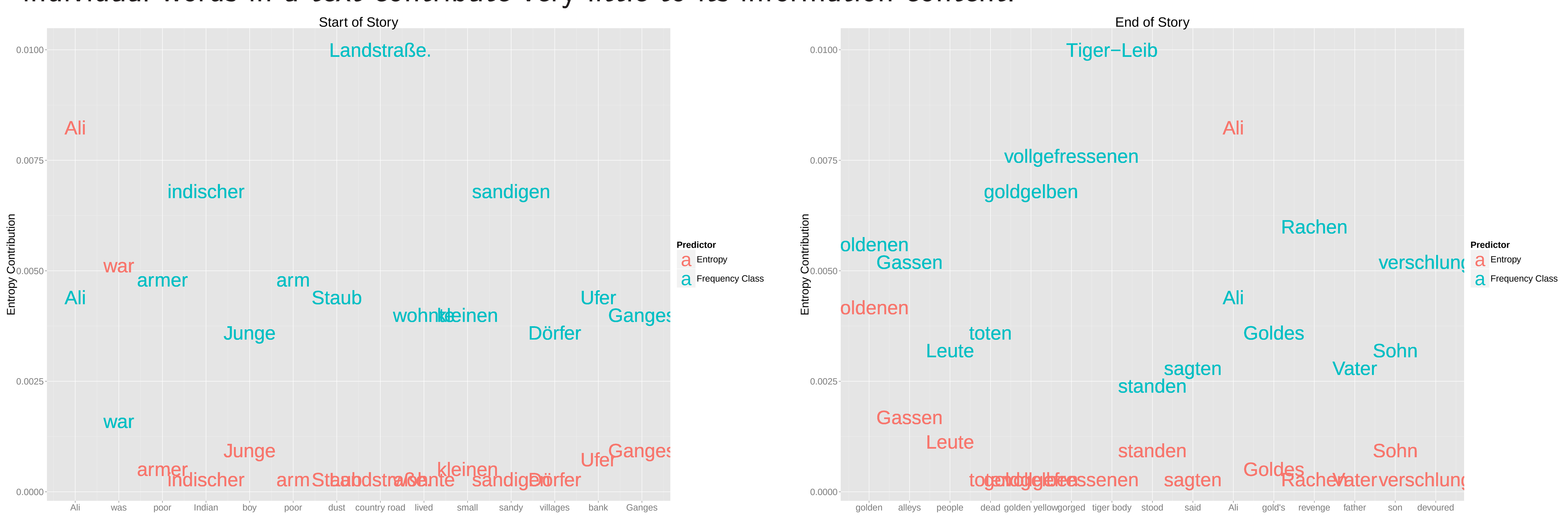
natural logarithm of length of utterance (scaling factor)

summed over all words in the utterance

As such entropy should provide a measure of expectation whose accuracy increases over the length of a text.

Entropy vs. Logarithmic Frequency Class

Locally calculated entropy and general word frequency show drastically different distributions. The majority of individual words in a text contribute very little to its information content.

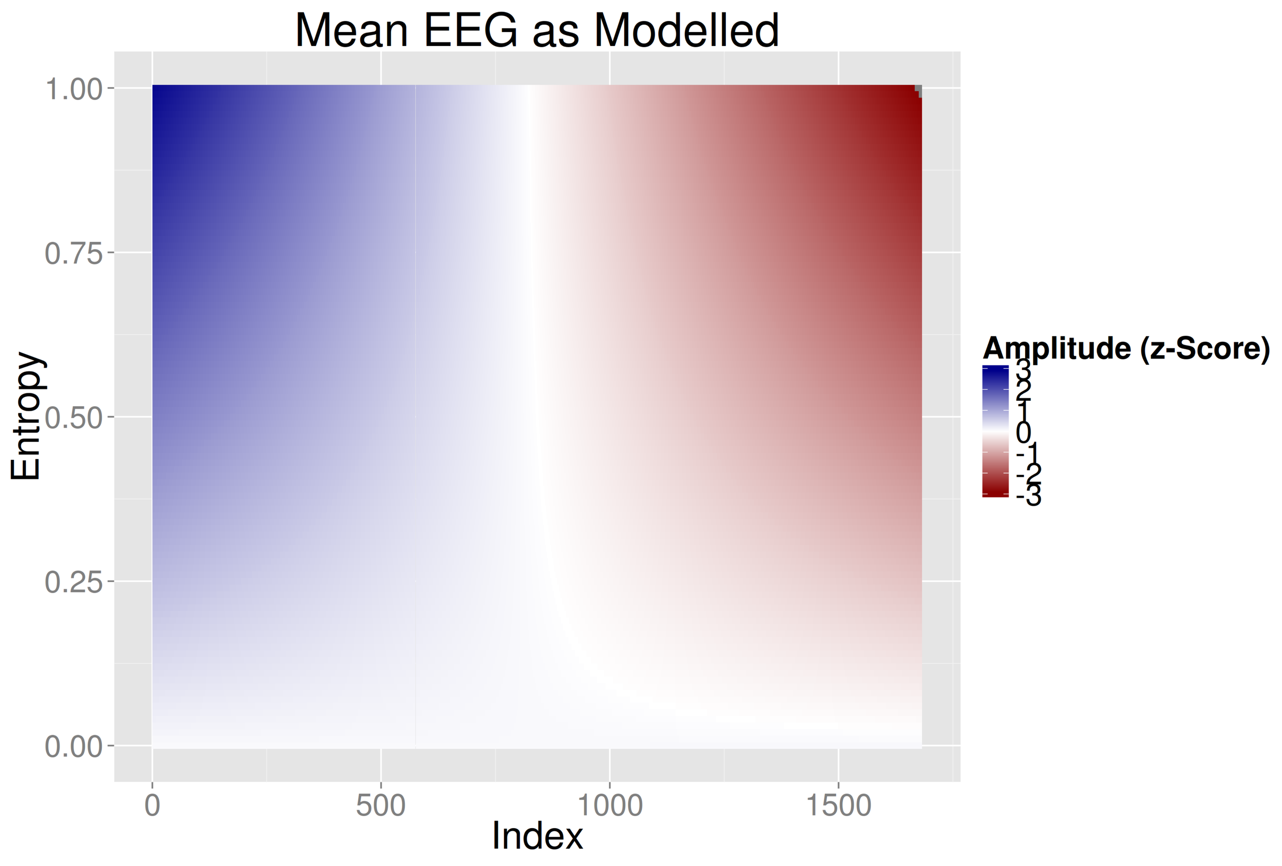


Frequency class was scaled by dividing by highest class times 100.

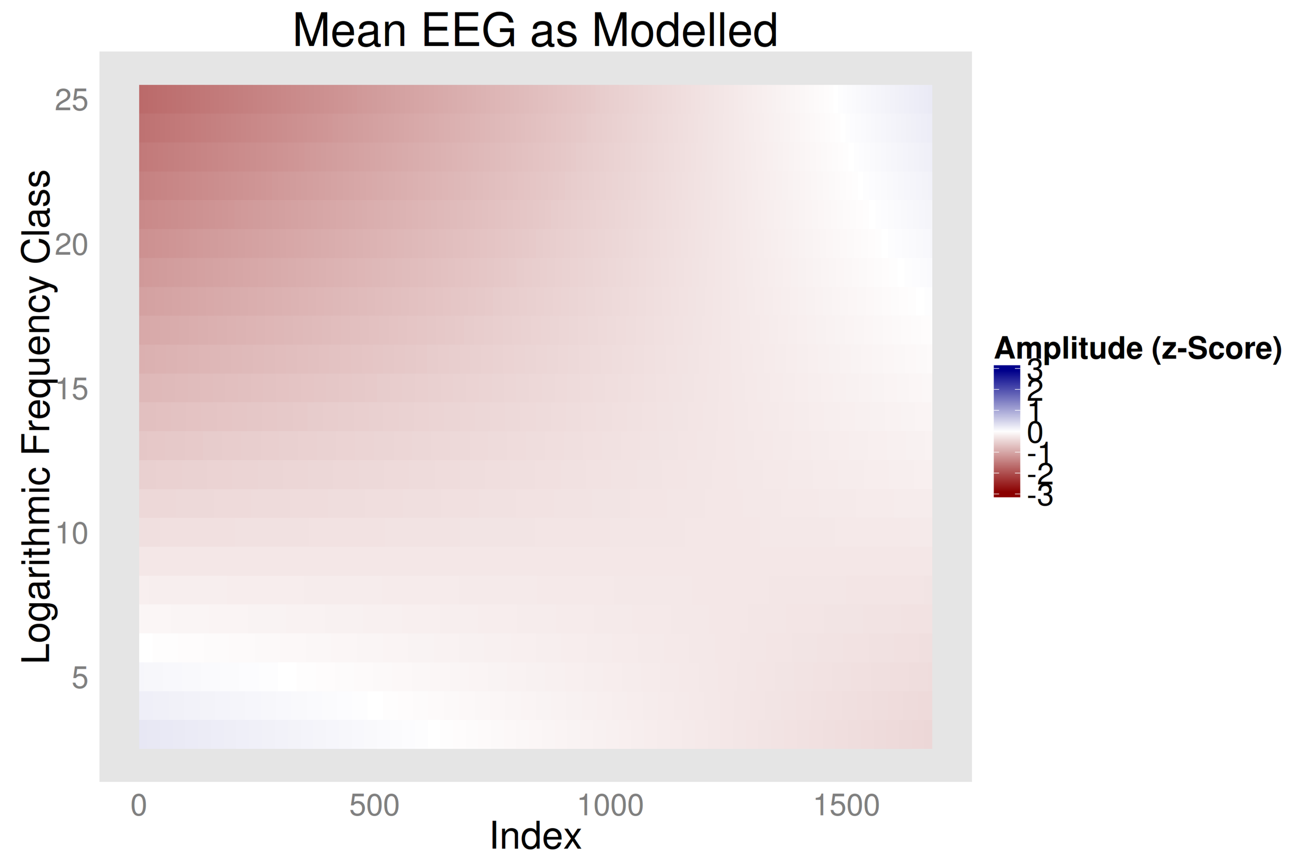
Experiment

52 subjects listened to a short story. 32-channel EEG was cleaned of artifacts with ICA [17] and 52*1682 segments extracted, time locked to the onset of content words.

Models



	Estimate	Std. Error	t value
index	0.00056	7.2e-05	7.8
entropy	3.1e+02	35	8.9
index:entropy	-0.39	0.038	-10
AIC:	557560	**logLik**: -278774	



	Estimate	Std. Error	t value
index	-0.00067	0.00015	-4.4
frequency	-0.092	0.01	-9.2
index:frequency	7.4e-05	1.1e-05	6.9
AIC:	616192	**logLik**: -308090	

Predictors were compared to single-trial mean EEG (Cz in the time window 300-500ms, typical for the N400 [18]) via mixed effects models using the R package `lme4`. Entropy interacts more strongly with position in the story than frequency does, reflecting the convergence of contextual expectation, and yields a much better model.

Comparison

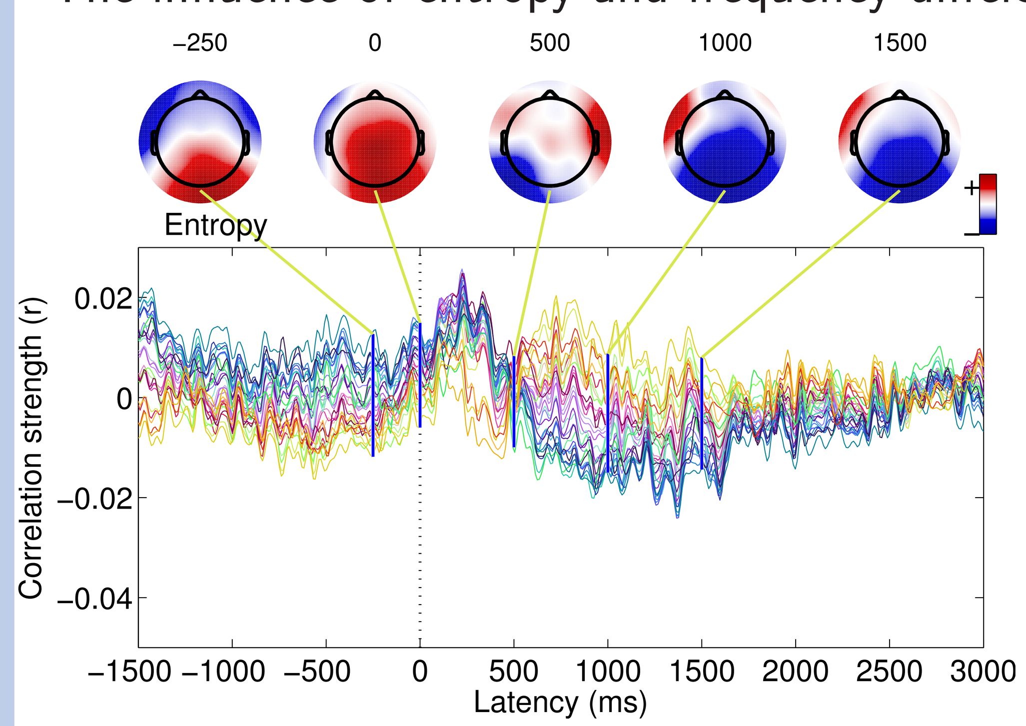
Moreover, entropy improves the frequency model much more than frequency improves the entropy model:

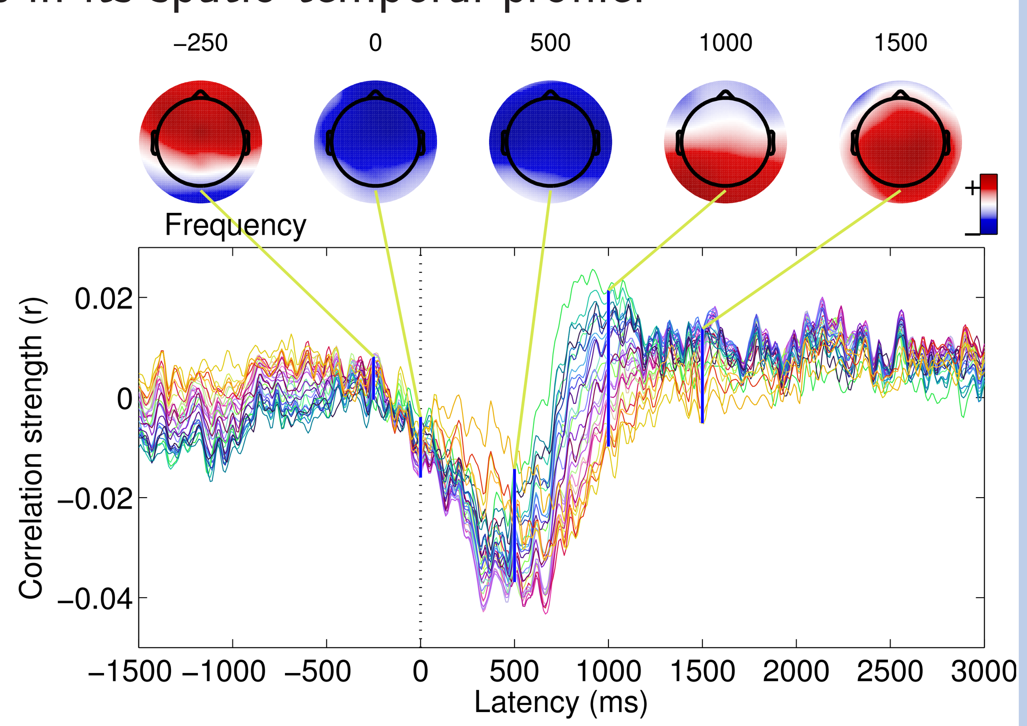
	Df	AIC	logLik	χ^2	χ^2-Df	$Pr(>\chi^2)$
frequency	6	616192	-308090			
with entropy	10	557123	-278551	59076	4	<0.001

	Df	AIC	logLik	χ^2	χ^2-Df	$Pr(>\chi^2)$
entropy	6	557560	-278774			
with frequency	10	557123	-278551	445	4	<0.001

Pointwise Correlation over the ERP

The influence of entropy and frequency differs in its spatio-temporal profile.





Conclusion

Hierarchical modelling of information theoretical measures allows for precise quantitative predictions. Brain activity indexes relative information content (cf. [1]).

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] C. E. Shannon (1948). *The Bell System Technical Journal.*

[6] C. E. Shannon (1951). *Bell System Technical Journal.*

[7] H. Akaike (1974). *IEEE Transactions on Automatic Control.*

[8] G. Schwarz (1978). *Annals of Statistics.*

[9] J. Hale (2003). *Journal of Psycholinguistic Research.*

[10] R. Levy (2008). *Cognition.*

[11] P. Garrard & B. Elvevåg (2014). *Cortex.*

[12] G. K. Zipf (1935). *The Psycho-Biology of Language.* Boston, MA: Houghton Mifflin Company.

[13] M. A. Montemurro (2014). *Cortex.* Language, Computers and Cognitive Neuroscience.

[14] C. Van Petten & M. Kutas (1990). *Memory and Cognition.*

[15] M. S. George, S. Mannes, et al. (1994). *Journal of Cognitive Neuroscience.*

[16] N. Lotze, S. Tune, et al. (2011). *Neuropsychologia.*

[17] I. Winkler, S. Haufe, et al. (2011). *Behavioral and Brain Functions.*

[18] M. M. Kutas & K. D. Federmeier (2011). *Annual review of psychology.*

[19] D. Bates, M. Maechler, et al. (2014). *lme4: Linear mixed-effects models using S4 classes.* R package version 1.1-7.