Semantic Update as a Predictor of Reading Time: Moving Beyond Word-Level Surprisal

Siddharth Gupta, Alessandro Lopopolo, Milena Rabovsky
Department of Psychology, University of Potsdam
siddharth.gupta@uni-potsdam.de

Surprisal has been ubiquitous in testing psycholinguistic hypotheses about lexical predictability and its effects on reading measures [1]. While surprisal captures predictability at the lexical-level, to predict per-word total fixation duration (gaze duration) we here test a complementary measure derived from an artificial neural network model of sentence comprehension that works at the level of sentence meaning: the Sentence Gestalt (SG) model [2]. This measure, called semantic update (SU), calculates how the introduction of each new word in a sentence updates the SG model's internal, predictive representation of sentence meaning [3, 4]. Mathematically SU, as reflected by the arrival of a new word w_n, is defined as: $SU_n = \sum_{i=1}^{k} |a_i(w_n) - a_i(w_{n-1})|$, where $a_i(w_n)$ is the activation at the i-th unit in the gestalt layer (with k units) as the network (see Fig 1 for the visualisation of encoder-decoder architecture in SG model) encounters the n-th word in the sentence. Hence, SU is the absolute error between the layer's activation before and after encountering a word. SU has previously been successfully applied to predict the N400 event-related potential (ERP) component [3, 4], the most widely used ERP component in research on language comprehension [5]. We hypothesize that reading behavior reflects both lexical-level prediction of individual word forms and the incremental integration of meaning at the sentence level - operationalised by word surprisal and SU respectively. As each word is encountered, it not only has a context-dependent predictability but also modifies the evolving representation of sentence meaning. We therefore expect that reading times reflect both how predictable a word is and how strongly it updates ongoing semantic representations.

In the current study, we examined the effect of SU as predictor of total word fixation duration (gaze duration) using an eye-tracking dataset of 46 native English-speaking participants who read 12 English texts [6]. To predict gaze duration, we fitted a linear mixed-effects model (lmer) that included SU, surprisal derived from a language model (LM) trained on the same dataset as the SG model, word length, word position, and word frequency, with random intercepts and slopes for SU and LM-based surprisal, grouped by both participant and word. We included LM-based surprisal [4] to assess whether the SU provides explanatory power beyond this information-theoretic predictor.

The results (Table 1) indicate that SU significantly predicts gaze duration ($\beta = 0.03, t = 2.81, p =< 0.001$) beyond the influence of LM-based surprisal, which as expected also has a significant impact ($\beta = 0.02, t = 2.27, p = 0.024$). Note-worthily, the effects of the predictors of interest are specific to content words as seen in lmer models run on the content and functional subsets as well as an interaction model with word category as the interaction term (Table 2). Besides, we also conducted an ANOVA comparing the first model to a model excluding SU and found that including SU significantly improves model fit $(p < 0.001)$.

The analysis reveals a significant relationship between gaze duration and the size of semantic updates. These findings suggest that gaze duration reflects the cognitive effort of updating semantic representations, emphasizing the role of sentence meaning-not just lexical-prediction in reading behavior.

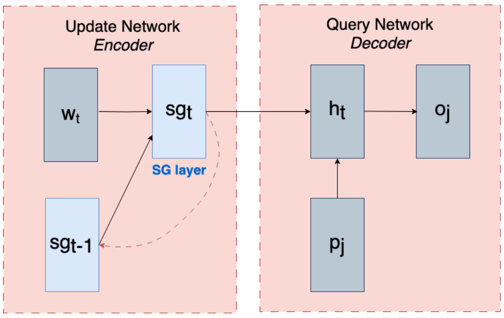

Figure 1: A visualisation of encoder-decoder architecture of Sentence Gestalt (SG) Model

	Estimate	CI	Std. Error	df	t value	Pr(>|t|)
Intercept	5.64	5.60 – 5.68	0.02	59.15	275.56	<**0.001**
Semantic Update	0.03	0.01 – 0.04	0.01	226.00	2.81	<**0.001**
LM Surprisal	0.02	0.00 – 0.04	0.01	191.46	2.27	**0.024**
Word Length	0.12	0.11 – 0.14	0.01	428.74	12.97	<**0.001**
Word Position	-0.02	-0.04 – -0.01	0.01	1612.57	-2.77	**0.005**
Word Frequency	-0.04	-0.06 – -0.01	0.01	458.16	-3.15	**0.002**

Table 1: Results of the linear mixed-effect model predicting gaze duration as a function of semantic update, LM-based surprisal, and word-level features.

	Estimate	CI	Std. Error	df	t value	Pr(>|t|)
(Intercept)	5.65	5.61 – 5.69	0.02	68.10	266.13	<**0.001**
Semantic Update	0.04	0.02 – 0.05	0.01	257.15	3.63	<**0.001**
LM Surprisal	0.04	0.01 – 0.06	0.01	357.83	2.94	**0.004**
Word Length	0.12	0.10 – 0.14	0.01	431.70	12.61	<**0.001**
Word Position	-0.02	-0.04 – -0.01	0.01	1626.01	-2.62	**0.009**
Word Frequency	-0.02	-0.05 – 0.00	0.01	511.86	-1.64	0.1
Word Category	-0.06	-0.10 – -0.01	0.02	601.67	-2.26	**0.024**
Semantic Update:Word Category	-0.05	-0.09 – -0.01	0.02	169.15	-2.52	**0.013**
LM Surprisal:Word Category	-0.03	-0.06 – -0.00	0.02	85.22	-2.02	**0.046**

Table 2: Results of the linear mixed-effect model predicting gaze duration as a function of semantic update, LM-based surprisal, word-level features and interaction term for word type (content vs functional) with reference level set to content words.

References

[1] Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*(3), 302–319. https://doi.org/10.1016/j.cognition.2013.02.013

[2] McClelland, J. L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes*, *4*(3-4), SI287–SI335. https://doi.org/10.1080/01690968908406371

[3] Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*, *2*(9), 693–705. https://doi.org/10.1038/s41562-018-0406-4

[4] Lopopolo, A., & Rabovsky, M. (2024). Tracking Lexical and Semantic Prediction Error Underlying the N400 Using Artificial Neural Network Models of Sentence Processing. *Neurobiology of Language*, *5*(1), 136–166. https://doi.org/10.1162/nol_a_00134

[5] Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology*, *62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

[6] Siegelman, N., Schroeder, S., Acartürk, C., Ahn, H.-D., Alexeeva, S., Amenta, S., Bertram, R., Bonandrini, R., Brysbaert, M., Chernova, D., Da Fonseca, S. M., Dirix, N., Duyck, W., Fella, A., Frost, R., Gattei, C. A., Kalaitzi, A., Kwon, N., Lõo, K., … Kuperman, V. (2022). Expanding horizons of cross-linguistic research on reading: The Multilingual Eye-movement Corpus (MECO). *Behavior Research Methods*, *54*(6), 2843–2863. https://doi.org/10.3758/s13428-021-01772-6