

Information density of encodings: The role of syntactic variation in comprehension



Les Sikos, Clayton Greenberg, Heiner Drenhaus and Matthew W. Crocker
Department of Language Science and Technology, Saarland University, Germany

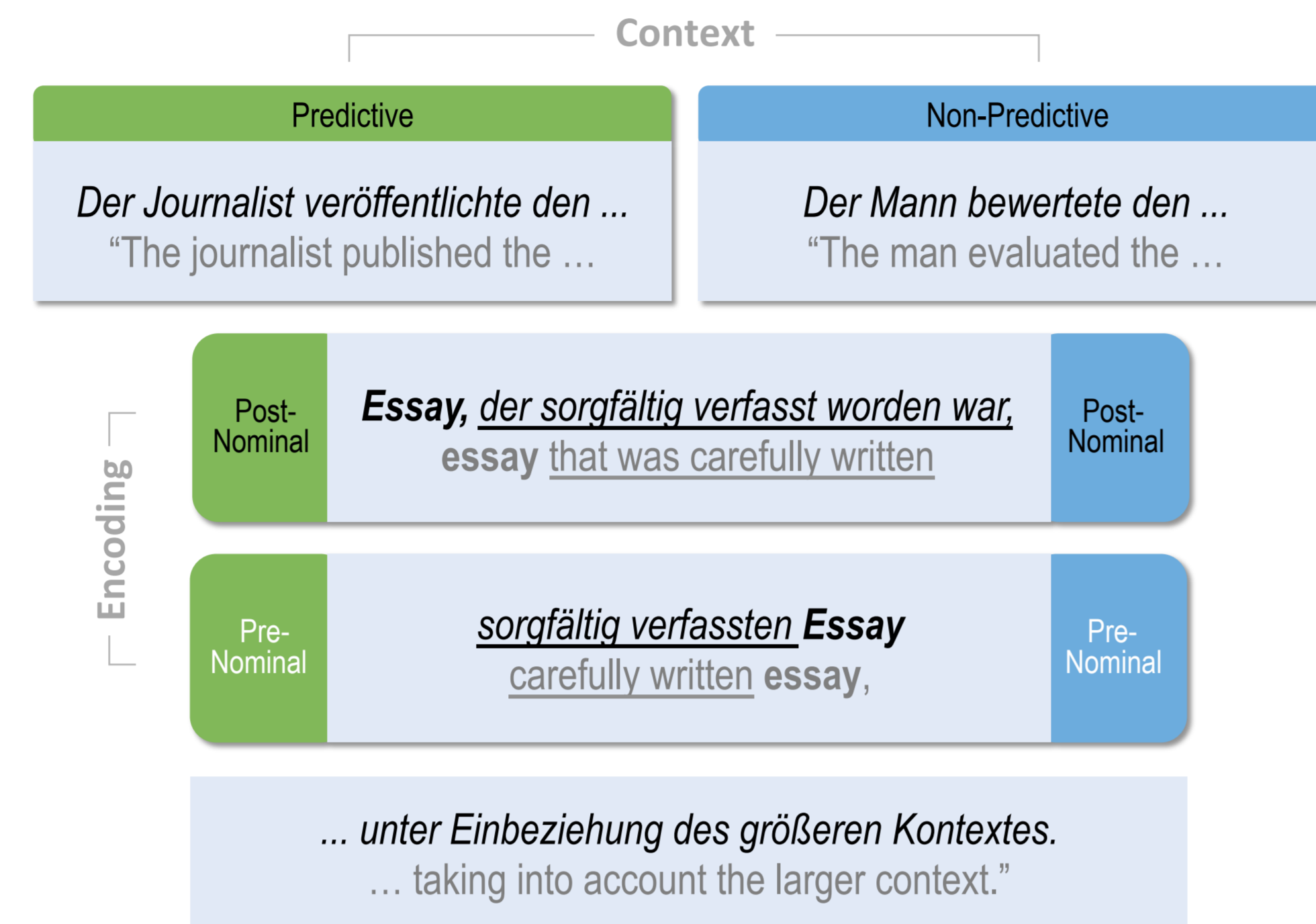


Introduction

- Uniform Information Density (UID) hypothesis links production strategies with comprehension processes
 - Predicts speakers utilize flexibility in encoding to increase uniformity in rate of information transmission, as measured by surprisal [1]
 - Evidence for UID comes largely from studies of word-level effects [1-4]
- **Goal:** Test whether comprehenders are sensitive during online processing to differences in information density of alternative syntactically-complex encodings

Materials

- 48 sets of German sentences crossing **Context × Encoding**
- Critical region: object NP
- RTs analyzed separately for object noun and modification region
- 48 Fillers with same structure but highly predictable object



Example Stimulus Set. Object noun (bold), modification region (underlined).

Methods

Cloze Probability and Contextual Constraint

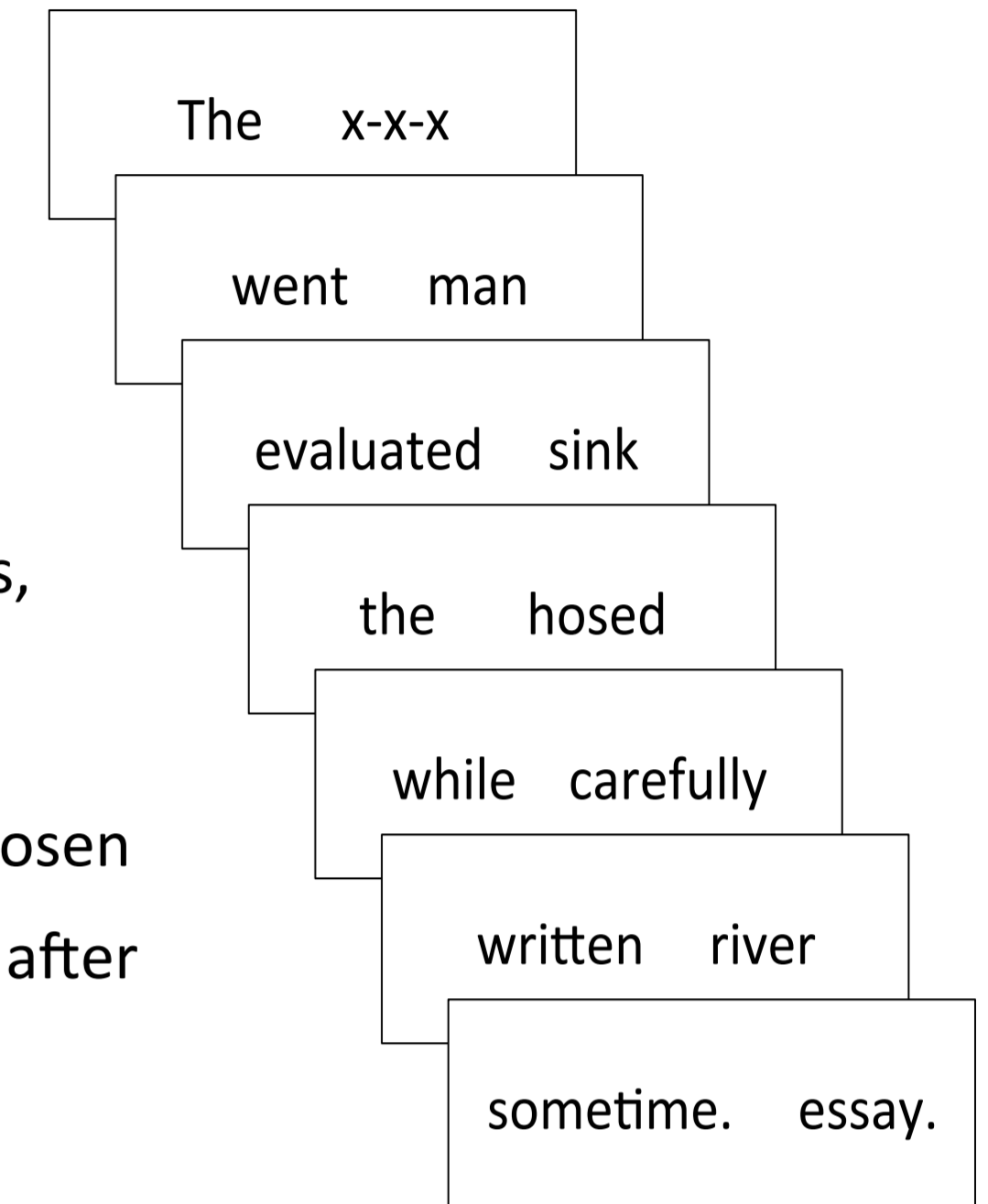
- Object nouns had higher cloze in predictive than non-predictive contexts, $p < .05$
- Contextual constraint greater for predictive than non-predictive contexts, $p < .001$

Language Model for computing Surprisals

- Interpolated modified Kneser-Ney 5-gram trained on German Wikipedia
- Final vocabulary size = 833,734
- 8:1:1 ratio; training = 666.5 M tokens
- Perplexities: training = 25, test = 201, stimuli = 1583
- Difference in perplexity between test and stimuli suggests German Wikipedia not ideal for stimuli

Participants

- 24 native German speakers
- Task** Grammaticality Maze
- Variation of SPR [5]
 - Less susceptible to spill-over [6]
 - Sentences presented word-by-word as sequence of forced choices between two alternatives, only one of which continues the sentence grammatically
 - Trial aborted if incorrect word chosen
 - Yes/No comprehension question after 1/3 of items



Results

Comprehension Q Accuracy

- $M = 97%$, confirming that participants were reading for meaning during G-maze task

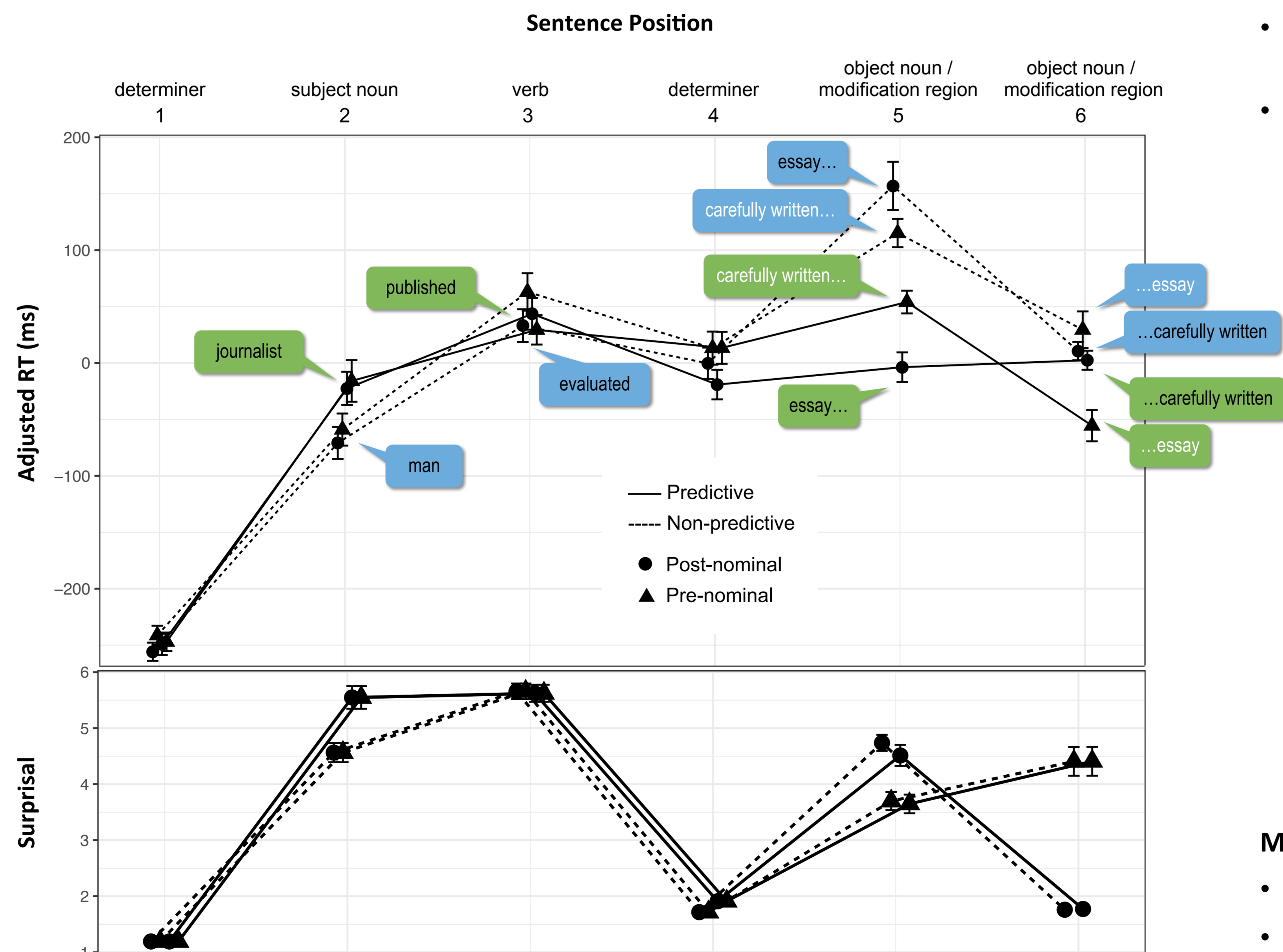
Response Time

- Analyses conducted on items completed through at least end of object NP ($M = 90%$)
- RTs were adjusted for word length and punctuation
- Adjusted RTs were analyzed using linear mixed models
- Because number of words modifying object nouns varied across items and conditions, adjusted RTs for modification region were averaged across all modifier words

Surprisal Profiles

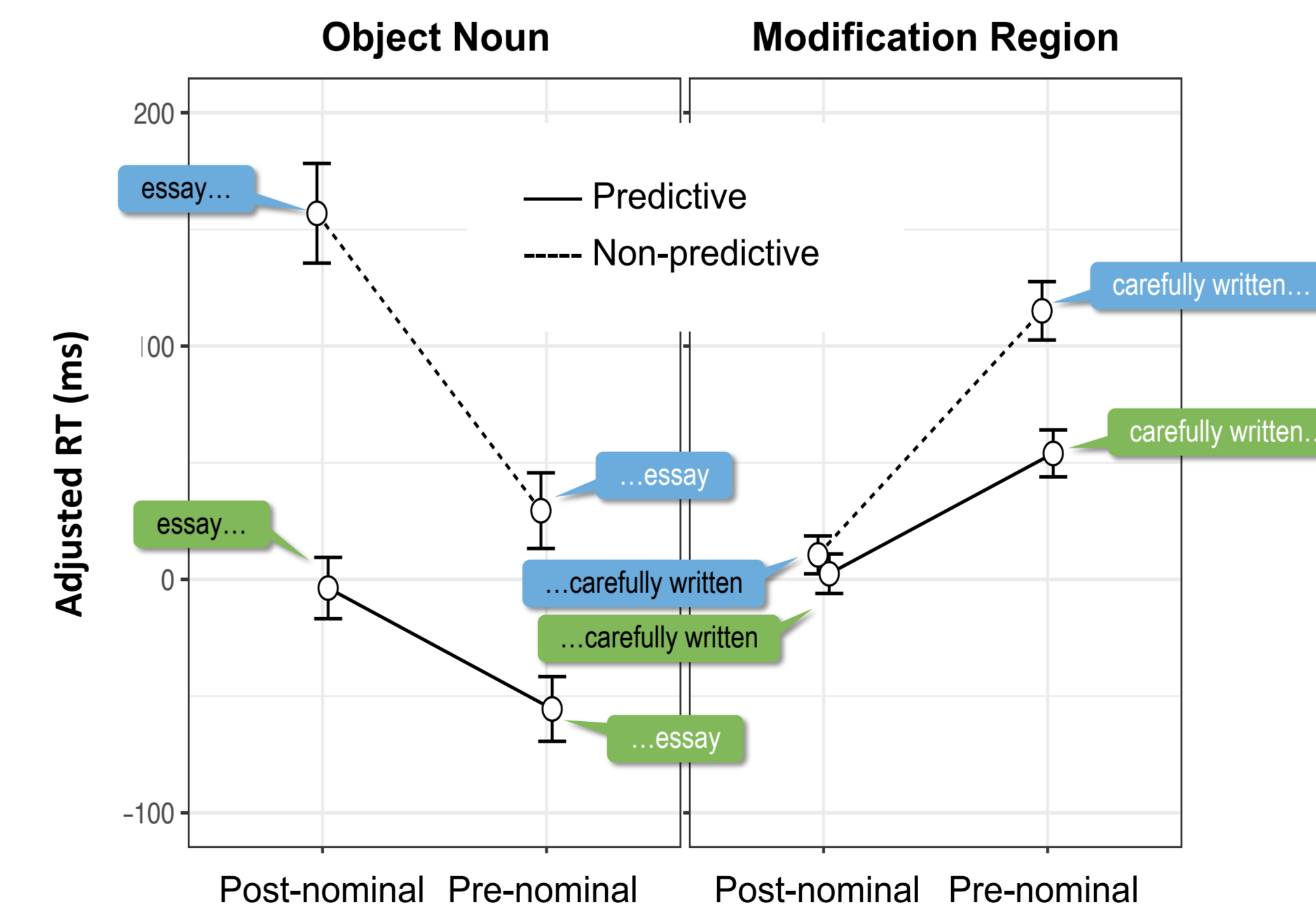
- Language model broadly confirmed assumption about distribution of surprisal across alternative encodings: pre-nominal was more uniform

Mean Adjusted Word-by-Word Reading Times per Condition



Object Noun Analysis

- Faster RTs in predictive than non-predictive conditions, $p < .001$
- Consistent with UID, in non-predictive conditions, RTs were faster for pre-nominally modified nouns than post-nominally, $p < .001$
- Pre-nominal facilitation effect was weaker for predictive conditions, resulting in Context × Encoding interaction, $p < .05$



Modification Region Analysis

- Encoding influenced modification processing in complementary way
- Pre-nominal modifiers read more slowly than post-nominal in both contexts, $p < .001$, but effect was greater in non-predictive, $p < .01$

Discussion

Results are consistent with UID

- RTs reduced for pre-nominally modified head nouns, but effect was larger in non-predictive than predictive contexts
- Comprehension effort and surprisal are similarly distributed across larger encodings
- Pre-nominal material has higher RT and surprisal than post-nominal material, but also modulates processing effort on head noun
- Additional influence of predictability on RTs suggests that readers expend more uniform effort across sentences when pre-nominals are used, especially in low predictability contexts

Next Steps

Currently investigating whether speakers are attentive to these effects on comprehenders when making encoding decisions during production

References

1. Jaeger (2010). Redundancy and reduction: speakers manage syntactic information density. *Cognitive Psychology*.
2. Frank & Jaeger (2008). Speaking rationally: Uniform information density as an optimal strategy for language production. *Cog Sci Proceedings*.
3. Levy & Jaeger (2007). Speakers optimize information density through syntactic reduction. In *Advances in Neural Info Processing Systems*.
4. Mahowald, Fedorenko, Piantadosi & Gibson (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*.
5. Forster, Guerrerera & Elliot (2009). The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods*.
6. Witzel, Witzel & Forster (2012). Comparisons of online reading paradigms. *Journal of Psycholinguistic Research*.