A Self-Paced Reading Benchmark Dataset of German Sentence Processing

Michael Vrazitulis & Shravan Vasishth

Introduction

- Benchmark data are an important tool for theory development and evaluating model predictions.
- The majority of benchmark data in sentence processing are limited to English (e.g., [6, 1]).

Our Work (in Progress)

- We collect self-paced reading benchmark data for a battery of postulated effects in German.
- So far, 216 out of target 1,100 Prolific subjects have been tested.
- 17 subjects are excluded due to chance-level accuracy on comprehension questions.
- We show the results so far, compared to qualitative and surprisal-based [2, 8] predictions.
- Bayes factors (BF_{10}) are used to evaluate the evidence for an effect being present.

Pre-Registration Protocol

osf.io/wpra9

Results on Spillover Region

osf.io/wtfhx

Predictions and Empirical Estimates

Included Experimental Designs

GPSD (2×2): Garden Paths From Subject-vs.-Direct-Object Ambiguity
Ambiguous/Unambiguous × S–O/O–S — closely replicating [10]

GPSI (2×2): Garden Paths From Subject-vs.-Indirect-Object Ambiguity
Ambiguous/Unambiguous × Active/Passive — loosely replicating [11]

AGAT (2×2): Agreement Attraction in Grammatical Sentences
Singular-/Plural-Controller × Match/Mismatch — closely replicating [4]

LOCO (2×2): Local Coherence
Coherent/Incoherent × Intervener/No-Intervener — closely replicating [12]

SBIN (2×2): Similarity-Based Interference
Subject-Cue [Yes/No] × Animacy-Cue [Yes/No] — closely replicating [13]

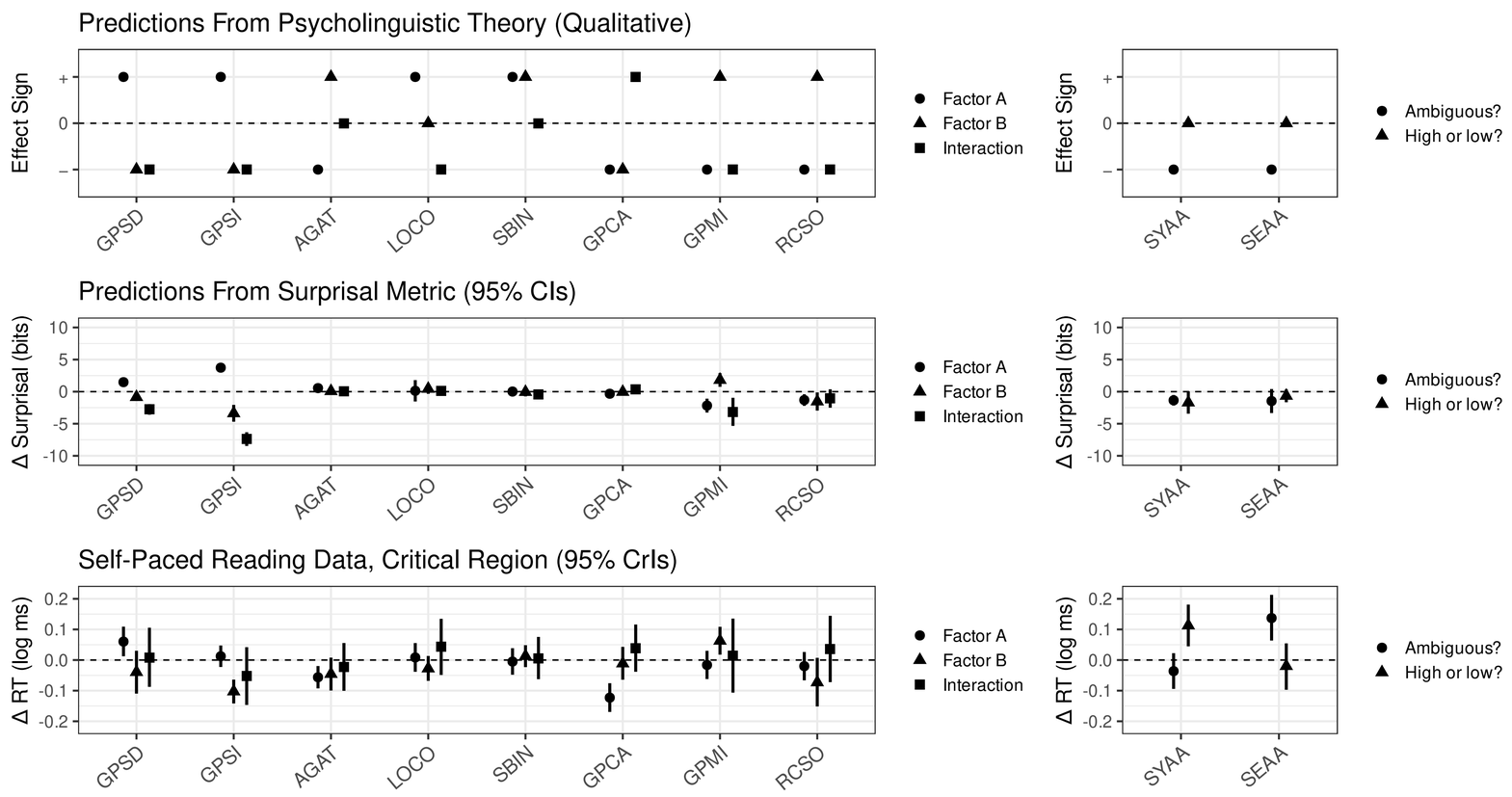
GPCA (2×2): Garden Paths From Coordination Ambiguity
NP-/VP-Coordination × AP-/PP-Modifier — closely replicating [7]

GPMI (2×2): Garden Paths From Modifier-vs.-Indirect-Object Ambiguity
Modifier/No-Modifier × Ambiguous/Unambiguous — closely replicating [3]

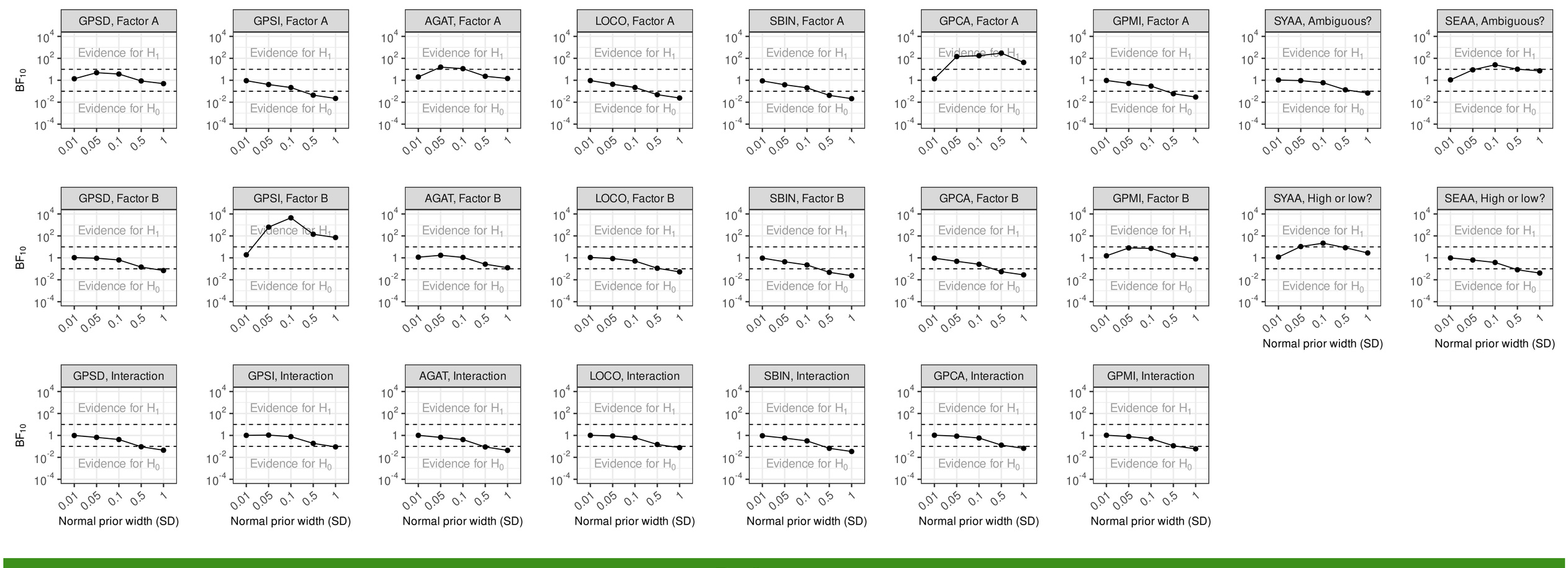
RCSO (2×2): Subject vs. Object Relative Clauses
Subject/Object × Double-/Single-Embedding — German adaptation of [5]

SYAA (3×1): Syntax-Based Attachment Ambiguity
High-/Low-/Ambiguous-Attachment — closely replicating [9]

SEAA (3×1): Semantics-Based Attachment Ambiguity
High-/Low-/Ambiguous-Attachment — German adaptation of [14]



Bayes Factor Analysis (Critical Region)



References

[1] R. Futrell et al. "The Natural Stories corpus: A reading-time corpus of English texts containing rare syntactic constructions". In: *Language Resources and Evaluation* 55 (2021), pp. 63–77. [2] J. T. Hale. "A probabilistic Earley parser as a psycholinguistic model". In: *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics*. Pittsburgh, PA, 2001. [3] J. Häussler. "Syntaktische und semantische Verarbeitungsprozesse bei der Analyse strukturell mehrdeutiger Verbfinalsätze im Deutschen: Eine empirische Untersuchung". PhD thesis. Free University of Berlin, 2001. [4] J. Häussler. "The emergence of attraction errors during sentence comprehension". PhD thesis. University of Konstanz, 2009. [5] F. Hsiao and E. Gibson. "Processing relative clauses in Chinese". In: *Cognition* 90.1 (2003), pp. 3–27. [6] K.-J. Huang et al. "Large-scale benchmark yields no evidence that language model surprisal explains syntactic disambiguation difficulty". In: *Journal of Memory and Language* 137 (2024), p. 104510. [7] L. Konieczny, B. Hemforth, and C. Scheepers. "Head position and clause boundary effects in reanalysis". In: *German Sentence Processing*. Ed. by B. Hemforth and L. Konieczny. Springer, 2000, pp. 247–278. [8] R. Levy. "Expectation-based syntactic comprehension". In: *Cognition* 106.3 (2008), pp. 1126–1177. [9] P. Logačev. "The role of underspecification in relative clause attachment: Speed–accuracy tradeoff evidence.". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 49.9 (2023), p. 1471. [10] M. Meng and M. Bader. "Mode of disambiguation and garden-path strength: An investigation of subject–object ambiguities in German". In: *Language and Speech* 43.1 (2000), pp. 43–74. [11] M. Meng and M. Bader. "Ungrammaticality detection and garden path strength: Evidence for serial parsing". In: *Language and Cognitive Processes* 15.6 (2000), pp. 615–666. [12] D. Paape and S. Vasishth. "Local coherence and preemptive digging-in effects in German". In: *Language and Speech* 59.3 (2016), pp. 387–403. [13] P. Schoknecht and S. Vasishth. "Do syntactic and semantic similarity lead to interference effects? Evidence from self-paced reading and event-related potentials using German". Under review. 2023. [14] M. J. Traxler, M. J. Pickering, and C. Clifton Jr. "Adjunct attachment is not a form of lexical ambiguity resolution". In: *Journal of Memory and Language* 39.4 (1998), pp. 558–592.