



# Implicit Learning and Syntactic Persistence: Surprisal and Cumulativity

T. Florian Jaeger<sup>1</sup> ([fjaeger@bcs.rochester.edu](mailto:fjaeger@bcs.rochester.edu)) & Neal E. Snider<sup>2</sup> ([snider@stanford.edu](mailto:snider@stanford.edu))

<sup>1</sup>Department of Brain & Cognitive Sciences, University of Rochester, <sup>2</sup>Department of Linguistics, Stanford University



## Abstract

We present evidence that syntactic persistence is linked to the updating and maintenance of probabilistic syntactic knowledge. We find that syntactic persistence is SURPRISAL-SENSITIVE and CUMULATIVE.

Our evidence comes from four corpus studies on conversational speech and one meta-analysis of production studies on priming.

## Cumulative Priming

- If syntactic persistence is due to implicit learning of syntactic structure, persistence should be long lasting. Evidence for this comes from lack of/slow decay of persistence effects in the absence of prime-target head lemma overlap (BockGriffin00; BockETAL07; HartsuikerETAL08).
- Another source of evidence for the longevity of persistence is cumulative priming, the cumulative of preceding primes beyond the closest prime. Evidence for cumulative priming so far comes from laboratory experiments only (KaschakETAL07; KaschakBorringer07).
- Evidence for cumulative persistence in spontaneous speech would provide further support for implicit learning accounts.

## Prime Surprisal

- We propose that the language processing system is set up in such a way that, whenever an instance of a structure is processed, it is seen as a piece of evidence that affects the structure's probability distribution
- Maintenance of probability distributions is assumed to be an inherent part of the language processing system.
- Less probable syntactic structures, if observed, lead to a bigger change in the probability distribution, which in turn leads to an increased probability of reusing the same structure (cf. HuberO'Reilly03 on perceptual persistence; HuberETAL01 on skill maintenance).
- More surprising primes should prime more strongly (i.e. lead to a bigger increase in the probability of repetition)

$$\text{surprisal}(X) = \frac{1}{\log(\text{probability}(X))}$$

- Surprisal-sensitive priming would also explain the well-known anti-frequency effect (the less frequent structure primes more, Bock86; V. Ferreira03, i.a.)

## Method of Analysis

### ① - ④ Mixed Logit Models

- Mixed logit models (*lmer*) in R 2.7.1, package *lme4* to predict speakers' structural choice in target utterances depending on a variety of priming related variables and other control predictors.
- Collinearity between predictors was removed from the model using centering and residuals (described as part of the studies)

### Data

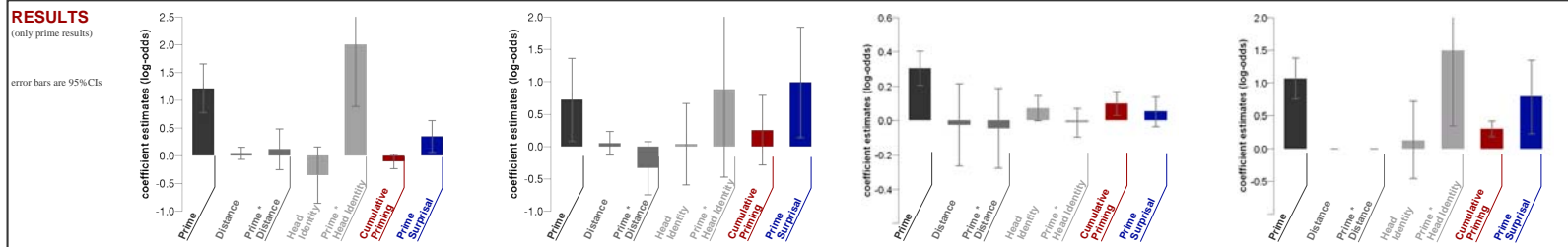
#### ① - ④ Paraphrase SWITCHBOARD (BresnanETAL02; MarcusETAL99)

- 642 telephone dialogues between 358 speakers of American English; about 800,000 words in 110,000 sentences.

#### ① Meta-analysis of laboratory experiment (KaschakETAL07)

- 6,375 prime and 2,193 target trials from 2 production experiments on cumulativity of syntactic persistence in ditransitives (276 subjects, 28 items)

	① Voice Alternation <i>The suspect was arrested. vs. They arrested the subject</i>	② Ditransitive Alternation <i>He gave the book to him. vs. He gave him the book.</i>	③ Optional <i>that</i> in CCs <i>Nobody believed (that) it would end this way.</i>	④ Ditransitive Alternation
<b>TARGETS &amp; PRIMES</b>	<b>Impersonal transit.:</b> 434 (20%) PASS vs. 1,702 (80%) ACT <b>w/ imp. tr. primes:</b> 221 (10%) PASS vs. 1,915 (90%) ACT	<b>Ditransitive targets:</b> 250 (22%) PO vs. 903 (78%) DO <b>w/ ditrans. primes:</b> 257 (22%) PO vs. 896 (78%) DO	<b>CC targets:</b> 1,155 (21%) THAT vs. 4,434 (79%) ZERO <b>w/ primes:</b> 969 (17%) THAT vs. 4,620 (83%) ZERO	<b>Ditransitive targets:</b> 653 (52%) PO vs. 597 (48%) DO <b>w/ ditrans. primes:</b> 673 (54%) PO vs. 577 (46%) DO
<b>PRIMING CONTROL PARAMS</b>	<b>Prime:</b> PAS (vs. ACT) <b>Prime-Target Distance &amp; interaction w/ Prime:</b> log-transformed distance in utterances <b>Head Identity &amp; interaction w/ Head Identity:</b> Prime & target verb lemma are identical (interaction effect beyond closest prime and head identity) <b>Cumulative Priming:</b> Empirical logit of cumulative proportion of PAS primes (beyond effect of closest prime) <b>Prime Surprisal:</b> Surprisal of closest PAS-prime given prime verb (beyond presence of closest prime and beyond target bias)	<b>Prime:</b> PO (vs. DO) <b>Prime-Target Distance &amp; interaction w/ Prime:</b> log-transformed distance in utterances <b>Head Identity &amp; interaction w/ Head Identity:</b> Prime & target verb lemma are identical (interaction effect beyond closest prime and head identity) <b>Cumulative Priming:</b> Empirical logit of cumulative proportion of PO trials preceding target (beyond effect of closest prime) <b>Prime Surprisal:</b> Surprisal of closest PO-prime given prime verb (beyond presence of closest prime and beyond target bias)	<b>Prime:</b> THAT (vs. ZERO) <b>Prime-Target Distance &amp; interaction w/ Prime:</b> log-transformed distance in words <b>Head Identity &amp; interaction w/ Head Identity:</b> Prime & target verb are identical (interaction effect beyond closest prime and head identity) <b>Cumulative Priming:</b> Cumulative count of CCs w/ THAT preceding target (beyond effect of closest prime) <b>Prime Surprisal:</b> Surprisal that closest CC prime has THAT given prime verb (beyond presence of closest prime and beyond target bias)	<b>Prime:</b> PO (vs. DO) <b>Prime-Target Distance &amp; interaction w/ Prime:</b> constant (always 3 trials) <b>Head Identity &amp; interaction w/ Head Identity:</b> Prime & target verb lemma are identical (interaction effect beyond closest prime and head identity) <b>Cumulative Priming:</b> Empirical logit of cumulative proportion of PO trials preceding target (beyond effect of closest prime) <b>Prime Surprisal:</b> Surprisal that closest prime is PO given prime verb (beyond presence of closest prime and beyond target bias)
<b>ADD. CONTROL PARAMS</b>	<b>Controls based on Snider08</b> <b>Target Bias:</b> Logit-transformed subcategorization bias of target verb as estimated from database <b>Accessibility of Patient in Target:</b> 3 parameters <b>Speaker Dialect:</b> Speaker from NYC area. (1 parameter) <b>Speaker:</b> as random effect	<b>Controls based on BresnanETAL07</b> <b>Target Bias:</b> Logit-transformed verb bias (calculated from BresnanETAL07) <b>Theme &amp; Recipient Accessibility:</b> 4 parameters each <b>Other controls parameters:</b> 2 parameters <b>Speaker:</b> as random effect	<b>Controls based on Jaeger06</b> <b>Fluency:</b> speechrate, disfluencies, etc. (4 parameters) <b>Accessibility of CC subject:</b> 5 parameters <b>Domain complexity:</b> 8 parameters <b>Other control parameters:</b> 2 parameters <b>Speaker:</b> as random effects	<b>Target Bias:</b> Logit-transformed subcategorization bias of target verb as estimated from a mixture of large corpora (Roland et al., 2008) <b>Speaker, Item, and Experiment:</b> as random effects  Items were counter-balanced in the original design of Kaschak and colleagues



SUMMARY	①	②	③	④
• No prime decay (contrary to ReiterETAL06; Szmrecsanyi05)	• All non-priming related controls had the expected effect (not displayed)	• All non-priming related controls had the expected effect (not displayed)	• All non-priming related controls had the expected effect (not displayed)	• All controls had the expected effect (not all displayed)
• Effect of Head Identity replicates	• Prime decay (the interaction of Prime and Prime-Target distance) did not reach significance, but approaches significance in the predicted direction.	• No prime decay and no significant effect of Head Identity	• No prime decay and no significant effect of Head Identity	• Clearly significant effects of Cumulative Priming (as shown in KaschakETAL07) and Prime Surprisal
• Significant effect of Prime Surprisal, but not of Cumulative Priming	• No significant effect of Head Identity	• Significant effect of Cumulative Priming, but Prime Surprisal does not reach significance.	• Significant effect of Cumulative Priming, but Prime Surprisal does not reach significance.	• Results do not change if non-ditransitive primes are included (total n=1,567)
• Additional studies not restricted to impersonal transitives, found an effect of cumulativity, but further studies are necessary.	• Significant effect of Prime Surprisal, but Cumulative Priming did not reach significance	• Identical effects were also found for optional <i>that</i> in non-subject-extracted relative clauses		

## Conclusions

As expected if speakers implicitly track probabilistic distributions of syntactic structures, and, more generally, as expected by implicit learning accounts, syntactic persistence seems to be sensitive to the surprisal of primes and long-lasting enough to be cumulative. The effects are, however, weak, so that further work is necessary.

## References

1. Bock, R. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18(3), 355-387. New York, NY: Wiley.

2. Bock, R., & Griffin, C. (2000). The persistence of syntactic priming: Transient activation or implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(1), 17-34.

3. Bock, R., & Griffin, C. (2001). Priming effects on language production: A probabilistic account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 1-16.

4. Bresnan, J., Cullen, J., Coombs, R., Neeb, M., Sauer, M., Warren, T., et al. (2007). Parameter analysis for improved generation. In *LREC 2007* (Edinburgh, Scotland).

5. Bresnan, J., Cullen, J., Neeb, M., & Sauer, M. (2005). Predicting the effect of the Prime. *Proceedings of the Royal Netherlands Academy of Science*.

6. Bresnan, J., & Chomsky, N. (2007). The minimalist syntactic tree. In *Minimalist syntax: The framework* (Ed. by R. N. Johnson, & M. Uriagereka). Cambridge, MA: MIT Press.

7. Kaschak, M., Boringer, S., Schwaninger, S., Spillmann, S., & Vanthienen, D. (2008). Syntactic priming persists while the lexical bias decays. *Language*, 84(2), 293-308.

8. Kaschak, M., Boringer, S., & Spillmann, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

9. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

10. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

11. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

12. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

13. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

14. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

15. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

16. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

17. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

18. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

19. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.

20. Kaschak, M., & Boringer, S. (2008). Persistence of syntactic priming and its interaction with lexical bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 1-16.