



## Event Complexity

Some verb phrases encode complex event structures that may introduce processing challenges, supported by prior work showing behavioral contrasts (arrows indicate longer RTs in original studies):

Study	Event Type	Stimuli
McKoon & Love 2011 (ML11)	Hit-type	The workmen <b>banged</b> the tiles.
	Break-type	The workmen <b>chipped</b> the tiles. ↑
Gennari & Poeppel 2003 (GP03)	State	The detective <b>disliked</b> his partner.
	Event	The detective <b>inspected</b> the crime scene. ↑
Brennan & Pytkäinen 2010 (BP10)	Object Experiencer	Atelic: Without a doubt, the child <b>scared</b> the precious kitten.
	Object Experiencer	Telic: Within minutes, the child <b>scared</b> the precious kitten.
	Subject Experiencer	Atelic: Without a doubt, the child <b>cherished</b> the precious kitten.
	Subject Experiencer	Telic: Within minutes, the child <b>cherished</b> the precious kitten. ↑

- With larger samples and online data collection, can we replicate the results from GP03, ML11, and BP10 in SPRT and SMS?
- Can we find similar results using the highly incremental maze task? Can it provide new insight regarding these processing costs?
- Where effects replicate, to what extent might they be explained by lexical prediction/surprisal?

Tasks	Replicated Paper	Participants
Self-Paced-Reading (SPRT)	ML11, GP03	75 native English speakers
Stop-Making-Sense (SMS)	ML11, GP03	131 native English speakers
Maze Task	ML11, GP03, BP10	48 native English speakers

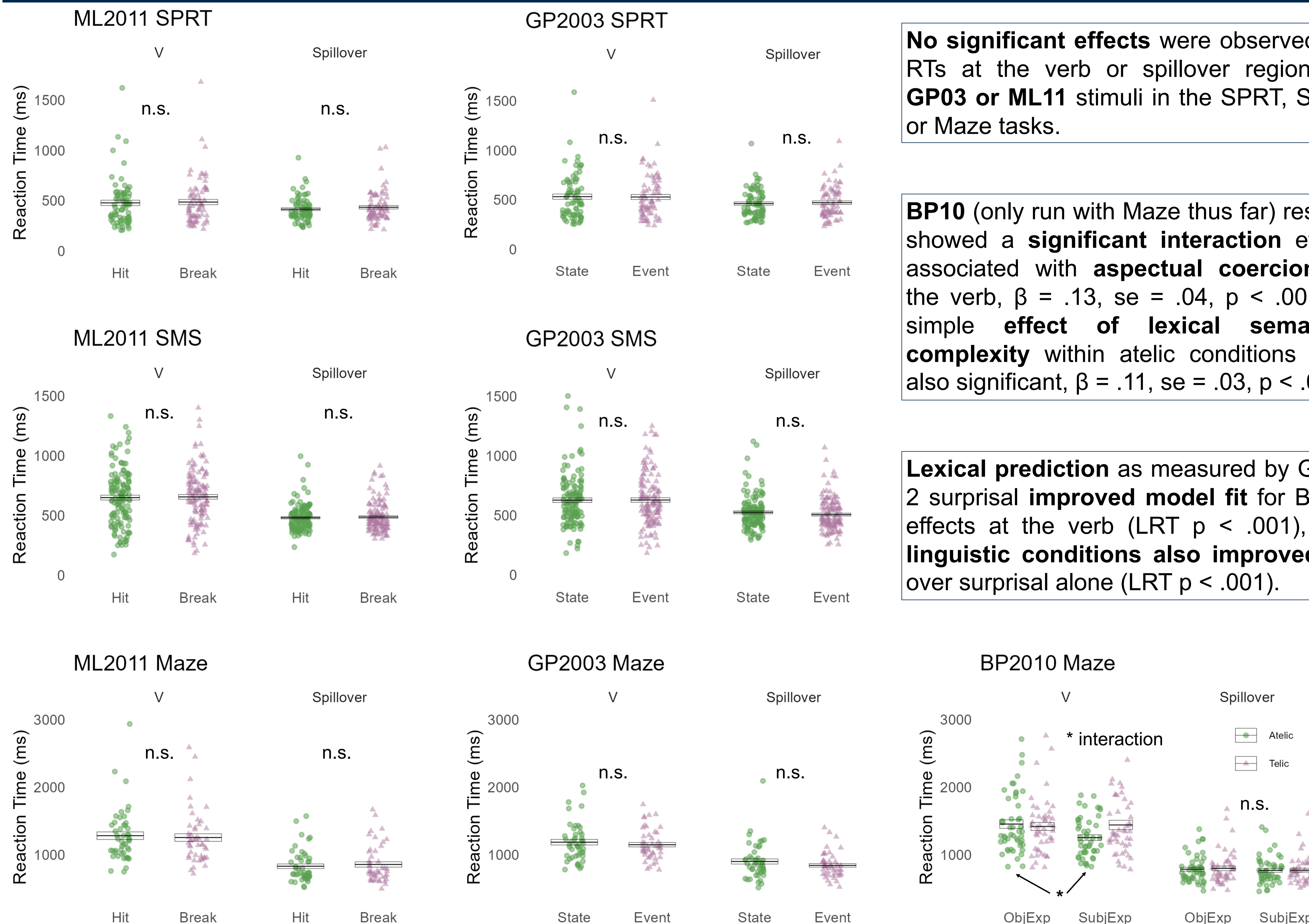
## Open Science

- Open sharing of stimuli, analyses, and datasets - key for future replication and furthering reproducible research.
  - Shared data facilitates meta-analyses, helping determine the **size and nature of effects** associated with specific linguistic variables
  - Transparent replication methods contribute to **cumulative evidence within sentence processing**.
- ML11, GP03, and BP10 were chosen as key studies on varied event complexity effects prior to rise of open science movement
- Conceptually replicating with larger online samples, multiple tasks, and updated statistical methods contributes to our broader understanding of the state-of-the-art in behavioral event processing.

## Acknowledgments

Thanks to Jared Diaz, Montana Thommes, and Sohee Chung for contributions to this work, and other members of the UM WordLab for valuable feedback.

## Results



No significant effects were observed for RTs at the verb or spillover region for GP03 or ML11 stimuli in the SPRT, SMS, or Maze tasks.

BP10 (only run with Maze thus far) results showed a **significant interaction** effect associated with **aspectual coercion** at the verb,  $\beta = .13$ ,  $se = .04$ ,  $p < .001$ . A simple **effect of lexical semantic complexity** within atelic conditions was also significant,  $\beta = .11$ ,  $se = .03$ ,  $p < .001$ .

Lexical prediction as measured by GPT-2 surprisal **improved model fit** for BP10 effects at the verb (LRT  $p < .001$ ), but **linguistic conditions also improved fit** over surprisal alone (LRT  $p < .001$ ).

## Analyses

- ML11 and GP03 are one-factor designs. BP10 is a 2x2 design crossing telicity and OE/SE verbs, predicting an interaction effect due to aspectual coercion in the telic + SE condition. BP10 also predicted a simple effect of “lexical semantic complexity” (causativity) for OE over SE within atelic conditions (modelled separately following original paper).
- Log RTs modelled as linear mixed effects with treatment-coded fixed effects for relevant linguistic conditions and word length, and random intercepts and slopes for participants and items.
- Where significant linguistic effects were observed, gpt2-medium surprisal was added as a fixed factor for model comparison.

## Discussion & Conclusions

- Results suggest **caution** in interpretation of prior results in behavioral studies on event structure and calls for further replication efforts.
- Although studies were conducted online, maze is very incremental task with high power and high accuracy even in internet-based studies.
- The replication of an aspectual coercion effect, but not the other event complexity effects suggests potentially different impacts on processing...
- However, both lexical semantic complexity (simple effect) in BP10 and ML11 described as “causativity” contrast – only BP10 effect replicated.
- Modelling of surprisal with the BP10 results also suggests syntactic or semantic complexity effects beyond lexical prediction.