Effects of foil processing, decision-making, and initial attention in the Maze task

John Duff[1], Pranav Anand[2], and Amanda Rysling[2]

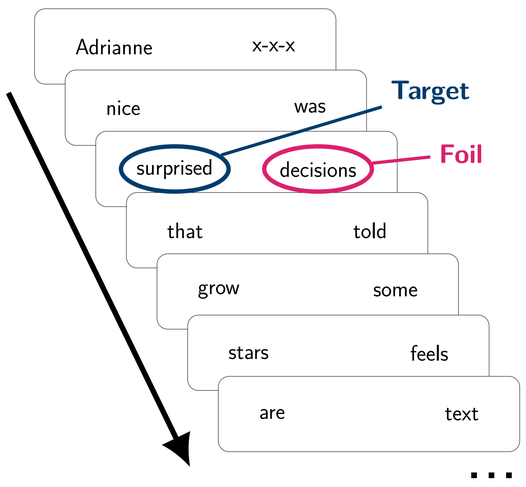
[1]Saarland University, Lang. Sci. & Tech. [2]UC Santa Cruz, Linguistics

jduff@lst.uni-saarland.de | AMLaP 30, 6 Sept. 2024

(PDF and materials)

Understanding the Maze task

- How do comprehenders approach Maze decisions?
- To what extent are measurements sensitive to Foil properties?

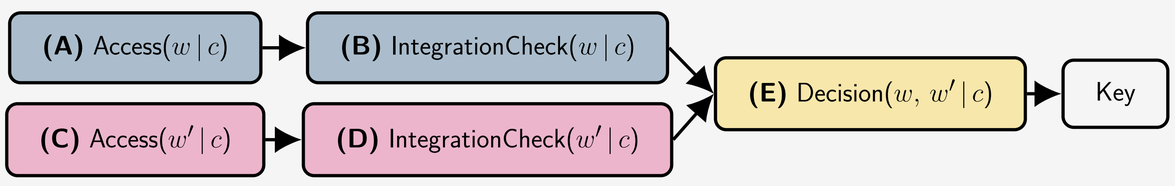


The Maze task [1, 5]

- Sequence of forced-choice decisions between suitable continuations (**Targets**) vs. inappropriate **Foils**.
- Foils often high-surprisal continuations from a language model [1].
- Choosing a Foil terminates the trial.
- Decision latencies for Target correlate with other reading measures for Target, without spill-over effects [1, 2, 5, 12].

A basic model of Maze latencies

Assume the latency of a correct Maze decision between Target w and Foil w' in context c is dependent on the duration of five processes:



- Target processing (A-B):
  - Only portion shared with typical reading processes
  - To use the Maze as a proxy for reading, experiments should aim to isolate differences here
- Foil processing (C-D):
  - Relative contributions are largely unknown, foils often controlled across conditions
- Decision-making (E):
  - By assumption, participants decide based on how easily each word integrates in c
  - Correct decisions would be slower when evidence is more balanced [7]
  - Initial evidence: Primed foils produce slower decisions [6]

Testing the model

Linking assumptions:

- Language model estimates of Surprisal($w \mid c$) are proportional to both difficulty of comprehending w [10] and the human plausibility of w as a continuation
- Access($w \mid c$), IntegrationCheck($w \mid c$) would be slower for higher Surprisal($w \mid c$)
- Decision(w, w') would be slower as Surprisal($w' \mid c$) − Surprisal($w \mid c$) approaches 0

Prediction: Effects of **Foil surp.** and **surp. diff.**, beyond effects of **Target surp.**

Methods

Data come from two large experiments in [4] using [1]'s method for high-surprisal Foils.

Surprisals for analysis were estimated (in bits) from GPT-2 [9] via `minicons` [8].

To make the strongest argument for the contribution of non-Target processes, we:

- fit initial, minimal regressions on odd trials only,
- used the minimal regression to extract residuals for even trials,
- then probed for the effect of other variables in regressions over those residuals.

Regressions were fit in `brms` [3] using weakly-constrained priors, centered linear predictors, and dummy-coded binary predictors.

❶
```
RT ~ Position * Trial + TargetSurp +
    (1 + Position * Trial + TargetSurp | Subject) + (1 | Item)
```

❷
```
Residual RT ~ (FoilSurp + SurpDiff) * TargetPosition * SideRepeat +
    (0 + (FoilSurp + SurpDiff) * TargetPosition * SideRepeat | Subject)
```

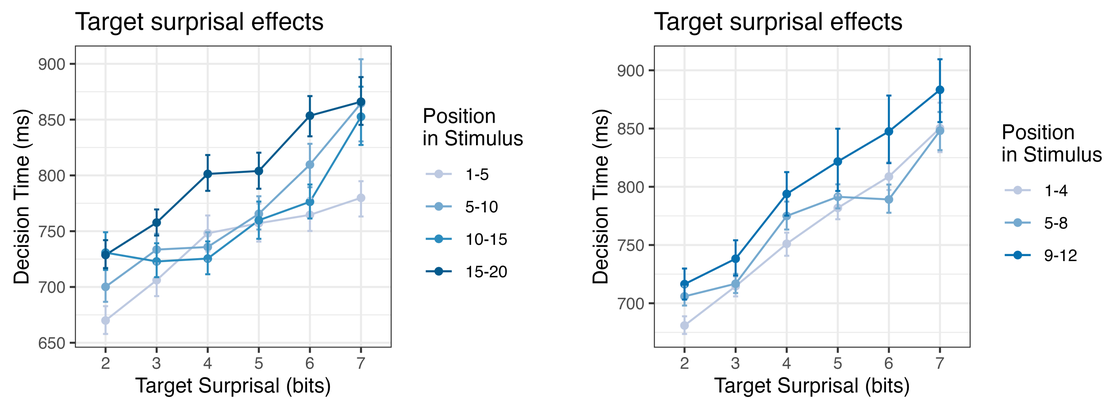
A more complex alternative (considering visual attention)

- Forced-choice ptcpts. sometimes make an absolute judgment on the first option [11]
- Mazers may pick a high-plausibility Target immediately if they attend it first
- Predicts: Foils matter less when the display drives Target-first reading, e.g.,
  - Target on L if participants default to read L-R, or
  - Target on side opposite previous Target if participants favor alternation

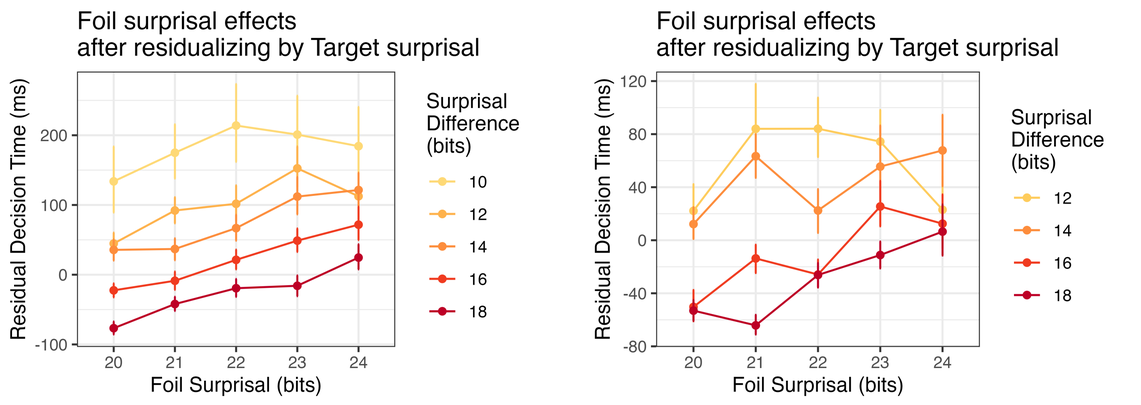
Results

Experiment A				Experiment B			
Ptcpts.	**Items**	**Words/Item**	**Correct RT Obs.**	**Ptcpts.**	**Items**	**Words/Item**	**Correct RT Obs.**
91	440	17–49 (med: 24)	180,309	143	336	5–26 (med: 21)	226,561

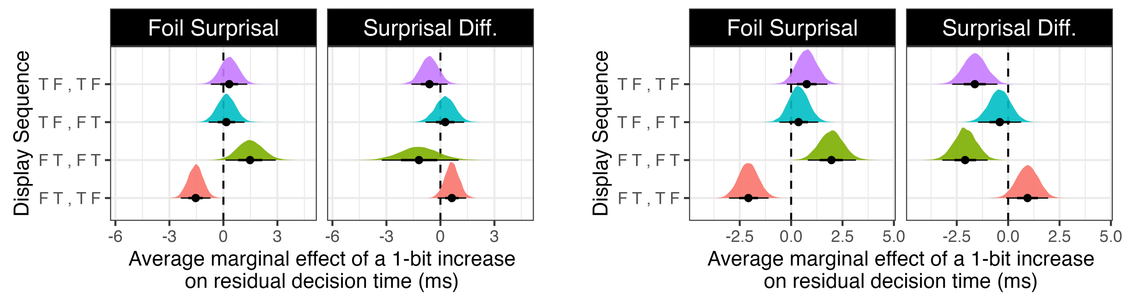
First analyses diagnosed credible **Target surprisal** effects ($\hat{\beta} = 10.08, 13.01$). Acceleration over the expt., but deceleration within the stimulus growing over the expt.



Second analyses diagnosed residual **(+) Foil surprisal and (−) difference effects**.



Foil effects were mediated by display, largest when Foil on L after Target on R (green) ($\hat{\beta}_{\text{Foil}} = 1.49, 1.96$; $\hat{\beta}_{\text{Diff}} = -1.18, -2.09$)



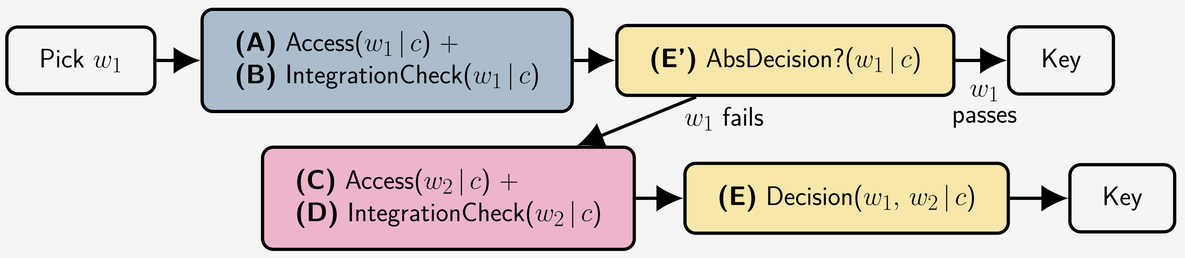
Discussion

Results validate effects of Foil processing and decision-making, but not on all trials.

Mazers appear to make absolute decisions [11] when they attend a good candidate first.

Display effects suggest Mazers distributed attention by L-R and/or alternation strategies.

A more complex model:



Consequences

Maze decision times for unexpected Targets may be inflated by:

- slow reading of appropriately unexpected Foils
- slow decision-making as Targets become harder to distinguish
- but less so in cases where displays allow Mazers to make absolute decisions

Accurate estimates of Target processing may require using Foil properties as covariates.

References:

[1] Boyce, Futrell & Levy. 2020. *J Mem Lang*.
[2] Boyce & Levy. 2023. *Glossa Psycholinguistics*.
[3] Bürkner. 2017. *J Stat Soft*.
[4] Duff. 2023. UC Santa Cruz Dissertation.
[5] Forster, Guerrera & Elliot. 2009. *Behav Res Meth*.
[6] Gallant & Libben. 2020. *The Mental Lexicon*.
[7] Luce. 1986. Oxford.
[8] Misra. 2022. ArXiv ms.
[9] Radford, Wu, Child, Luan & al. 2019. OpenAI.
[10] Shain, Meister, Pimentel, Cotterell & Levy. 2024. *PNAS*.
[11] Starns, Chen & Staub. 2017. *J Mem Lang*.
[12] Witzel, Witzel & Forster. 2012. *J Psycholing Res*.

Acknowledgements: Many thanks to RAs Sebastian Bissiri and Kasey La, and further thanks to Iza Škrjanec, Marian Marchal, Merel Scholman, Katja Häuser, and Vera Demberg for helpful discussions and feedback on this work.