

## Task influences on lexical underspecification: Insights from the Maze and SPR

John Duff, Adrian Brasoveanu, and Amanda Rysling (UC Santa Cruz)

In assigning an incremental interpretation to linguistic input, sometimes a resolution to a temporary ambiguity will prove incompatible with downstream material, leading to costly reanalysis. However, in some cases comprehenders appear to avoid this cost. One approach to such observations is to posit that comprehenders can **underspecify** some input, delaying full commitment and interpretation [2,3,7].

One set of data often taken as evidence for underspecification is [5]'s landmark eyetracking investigation of polysemy, words with multiple meanings which share core features (e.g. *newspaper* as printed object vs. corporate entity) and homonymy, properly ambiguous words with entirely distinct meanings (e.g. *jam* as fruit spread vs. blockage). [5] report that when disambiguation to a less frequent meaning follows a polyseme, reanalysis cost (first-pass RT, probability of regressions out) is less than when such disambiguation follows a homonym. This difference has been argued to follow from an account where decisions among meanings are initially underspecified only when choosing between overlapping meanings [5-7].

From findings of this sort, however, it's unclear why polyseme specification is delayed. We might consider two hypotheses: first, that underspecification is **utility-based** (effective under typical comprehension strategies); or, underspecification may be **necessary** due to some property of lexical representation and semantic commitment. The present study seeks to address this question by replicating [5] across additional tasks: if underspecification is deployed strategically, it will be sensitive to changes in a participant's priorities.

In the first experiment, we use [4]'s Maze, in which participants advance word-by-word by making decisions between the correct continuation of a sentence and a foil (see Fig. 1). In particular, we will use the A-Maze of [1], where foils are words with high surprisal in the existing context. If a participant chooses a foil instead of a target, the trial terminates.

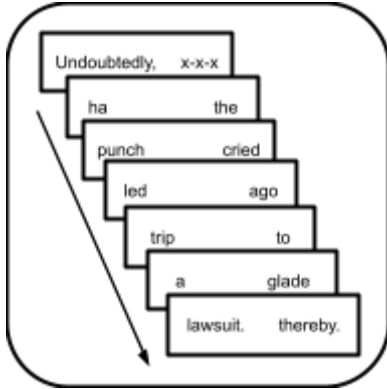
Participants in the Maze must engage in eager interpretation to maximize their ability to proceed through the stimulus, making underspecification a less useful strategy than in natural reading. A utility-based account of underspecification then predicts that in the Maze, polysemes may exhibit reanalysis costs similar to homonyms. Alternately, a hypothesis under which underspecification is necessary predicts we should replicate [5], with some interaction in RTs (here, response latencies) for the disambiguation region, such that late disambiguation shows greater costs for homonymy.

**Expt. 1** ( $n=48$ ) presented two sets of 32 items featuring polysemy (1) and homonymy (2) in the Maze. Each set crossed disambiguation POSITION (EARLY/LATE) X MEANING (M1/M2), after [5], with dominance established by acceptability norming. Note that EARLY conditions feature cataphoric dependencies. Participants saw items Latin-squared and randomized with 128 fillers.

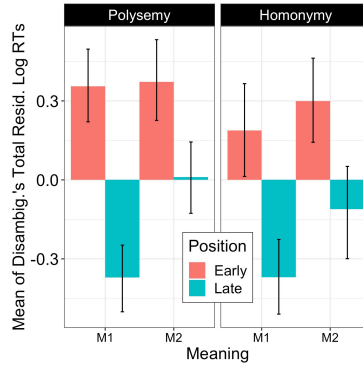
Log RTs in the disambiguator, residualized over position and length, were analyzed in a Bayesian-fit linear mixed-effects model (Table 1). We observe a POS main effect we link to the lack of cataphora in LATE, a POS X MEANING interaction indicative of a cost for late disambiguation to M2 for polysemes, and no interaction terms suggesting a difference for homonyms.

**Expt. 2** ( $n=48$ ) presented the same items in fixed-window SPR for minimal comparison with the Maze. Analysis reveals a POS X TARGET interaction such that late disambiguation bears larger costs for homonyms than polysemes, consistent with [5]'s underspecification findings.

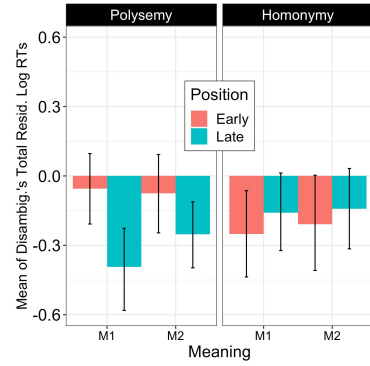
An account where underspecification is necessary makes the wrong predictions. Instead, we observe task-dependent variation in line with the utility account: polyseme underspecification is avoided in the Maze, a fact we attribute to the task's demands for eager interpretation. Researchers using the Maze should be aware that it imposes unique task pressures, but awareness of this fact can allow us to see where standard online behavior derives from strategic deployment of the language processing architecture rather than its limitations or requirements.



**Fig. 1.** A depiction of a toy A-Maze trial.



**Fig. 2.** Mean total residualized log RTs in the disambiguating region (E1).



**Fig. 3.** Mean total residualized log RTs in the disambiguating region (E2).

Fixed Effect	Expt. 1 (Maze)				Expt. 2 (SPR)			
	Mean	SD	95% CI		Mean	SD	95% CI	
POSITION: LATE	-0.74	0.12	-0.97	-0.50 *	-0.33	0.13	-0.58	-0.77 *
TARGET: HOMONYMY	-0.17	0.15	-0.46	0.12	-0.20	0.13	-0.45	0.06
POSITION x MEANING	0.35	0.16	0.05	0.66 *	0.16	0.18	-0.19	0.50
POSITION x TARGET	0.16	0.17	-0.18	0.49	0.44	0.18	0.09	0.79 *
POS x M x TARGET	-0.16	0.23	-0.61	0.29	-0.19	0.25	-0.67	0.30

**Table 1.** Excerpted mixed-effects models fit to total resid. log RTs in disambiguating region.

(1) **POLYSEMY** (disambiguating region)

- Unfortunately, after it was soaked with rain the newspaper was destroyed. [EARLY,M1]
- Unfortunately, after it lost its advertising profits the newspaper was destroyed. [E.,M2]  
(x-x-x intend in job lips discover obtain kid conducted add extension.)
- Unfortunately, the newspaper was destroyed after it was soaked with rain. [LATE,M1]
- Unfortunately, the newspaper was destroyed after it lost its advertising profits. [L.,M2]  
(x-x-x kid conducted add extension intend in job lips discover obtain.)

(2) **HOMONYMY** (disambiguating region)

- Reportedly, after it made his toast soggy the jam displeased Tom. [EARLY,M1]
- Reportedly, after it doubled his morning commute the jam displeased Tom. [EARLY,M2]  
(x-x-x, come fit detail sir thinks begin kept ours indecision Need.)
- Reportedly, the jam displeased Tom after it made his toast soggy. [LATE,M1]
- Reportedly, the jam displeased Tom after it doubled his morning commute. [LATE,M2]  
(x-x-x, kept ours indecision Need come fit detail sir thinks begin.)

**References**

- [1] Boyce, V., Futrell, R., & Levy, R. P. 2020. *Journal of Memory and Language* 111: 1-13.
- [2] Egg, M. 2010. *Language and Linguistics Compass* 4: 166-181.
- [3] Ferreira, F., & Patson, N. D. 2007. *Language and Linguistics Compass* 1: 71-83.
- [4] Forster, K. I., Guerrera, C., & Elliot, L. 2009. *Behavior Research Methods* 41: 163-171.
- [5] Frazier, L., & Rayner, K. 1990. *Journal of Memory and Language* 29: 181-200.
- [6] Frisson, S., & Frazier, L. 2004. Poster at CUNY 17, University of Maryland.
- [7] Frisson, S. 2009. *Language and Linguistics Compass* 3: 111-127.