Semantic commitment and lexical underspecification in the Maze

John Duff, Adrian Brasoveanu, and Amanda Rysling (UC Santa Cruz Linguistics)
jduff@ucsc.edu, abrsvn@ucsc.edu, rysling@ucsc.edu

A complete model of incremental sentence processing must encompass how and when lexical ambiguities are resolved in real time. In a landmark study, [5] demonstrated in eyetracking while reading that such resolution depends on the nature of the ambiguity, contrasting homonymy and polysemy. In homonymy, one lexical form has multiple non-overlapping meanings, e.g. the noun *jam* as fruit spread or traffic blockage. In polysemy, the multiple meanings are related by a core set of features, e.g. *newspaper* as printed object or corporate entity. [5] report that measures of reanalysis cost (increased first-pass RT and probability of regressions out of the disambiguating region) are greater for homonymy than polysemy when disambiguation to a less frequent meaning occurs after the ambiguous form.

Today, this finding is part of a body of evidence for the flexible **underspecification** of polysemy [7], echoing similar proposals in other domains [2,3]. It appears that the underspecification of polysemous forms persists only to the sentence boundary [6], but it remains an open question why this is. 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.

The present study seeks to replicate [5] for homonymy and polysemy with a different reading measure, the Maze task [4]. In this task, 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.

The Maze requires strictly incremental comprehension at each decision point [4], and therefore enables us to investigate how susceptible underspecification of polysemy is to task pressures. The utility-based account of underspecification predicts that here, polysemes may be incidentally specified immediately, like homonyms. Alternately, a hypothesis under which underspecification is necessary predicts we should replicate [5], as a 2nd-order diff. of diffs. in RTs (here, response latencies) for the disambiguation region, such that late disambiguation to a less frequent meaning prompts slower RTs more with homonymy than polysemy.

Our experiment (*n***=48)** used two sets of 32 items featuring polysemy (1) and homonymy (2), each crossing DISAMBIGUATION POSITION (EARLY/LATE) x MEANING (M1/M2), after [5]. All homonymy targets featured two inanimate nominal meanings, and polysemy targets featured concrete and abstract meanings. Participants saw items Latin-squared and randomized with 128 fillers. Foils were generated using the method in [1] (same foils across conditions).

Analysis was conducted over log RTs in the disambiguator, residualized over position and length, via a Bayesian-fit linear mixed-effects model with maximal appropriate random effects (Table 1). While we observe a significant interaction of POS x MEANING indicative of a cost for late disambiguation to M2, we fail to observe that this cost is stronger for homonyms, thus mismatching the predictions of the hypothesis that underspecification is necessary.

Results from this study thus suggest that underspecification of polysemy cannot be an architectural necessity. Indeed, if Maze task participants have a reason to commit as early as possible, then we might conclude underspecification is instead an optional heuristic, foregone here due to atypical task pressures. Nevertheless, at this stage one might worry whether the critical effect of [5] can be replicated at all. To more fully investigate our account, we plan to ultimately compare across tasks (SPR and, eventually, eyetracking) and across multiple constructions purported to exhibit underspecification. If the immediate specification observed here can indeed be attributed to the nature of incremental semantic commitment in the Maze in particular, the task may be a powerful tool to test the limits of the online comprehension of ambiguous input.

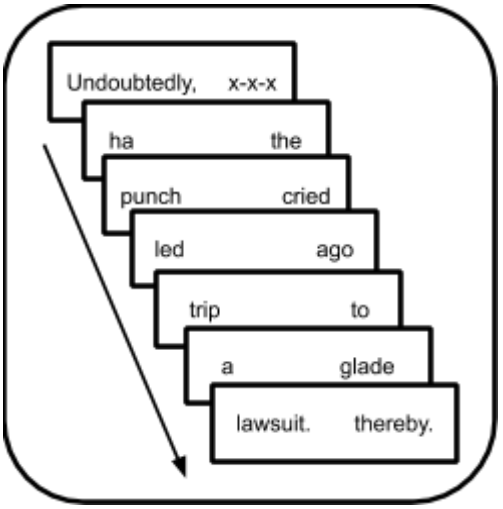

Fig. 1. A depiction of a toy A-Maze trial.

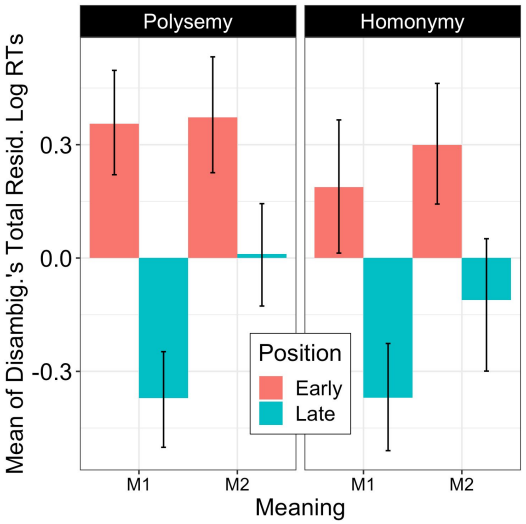

Fig. 2. Mean total residualized log RTs in the disambiguating region by condition.

Fixed Effect	Mean	SD	95% CI		
POSITION: LATE	-0.74	0.12	-0.97	-0.50	*
TARGET: HOMONYMY	-0.17	0.15	-0.46	0.12	
POSITION x MEANING	0.35	0.16	0.05	0.66	*
POS x M x TARGET	-0.16	0.23	-0.61	0.29	

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

(1) **POLYSEMY** (disambiguating region)
a. Unfortunately, after it was soaked with rain the newspaper was destroyed. [EARLY,M1]
b. 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.*)
c. Unfortunately, the newspaper was destroyed after it was soaked with rain. [LATE,M1]
d. 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)
a. Reportedly, after it made his toast soggy the jam displeased Tom. [EARLY,M1]
b. 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.*)
c. Reportedly, the jam displeased Tom after it made his toast soggy. [LATE,M1]
d. 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.
[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.