Semantic commitment and lexical underspecification in the Maze

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

Delaying Commitment

Does the processor underspecify out of utility or necessity?

Two types of lexical meaning multiplicity

Linguists distinguish two classes of lexical meaning multiplicity (loosely called 'ambiguity'), where a word has multiple meanings.

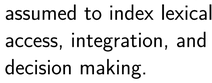
Polysemy: meanings are related by core features, e.g. *newspaper* as printed object or corporate entity.

Homonomy: (ambiguity in a strict sense) meanings are fully distinct, e.g. *jam* as fruit spread or traffic blockage.

In eyetracking, sentences where homonyms are disambiguated late to a less common meaning exhibit more signs of costly reanalysis than polysemes, e.g. higher probability of regressions [1].

(1) Unfortunately, the newspaper was destroyed after it lost its advertising profits.

(2) Reportedly, the jam displeased Tom after it doubled his morning commute.



Later work [2, 3] has framed this as evidence for the online **underspecification** of polysemy, *pace* [4, 5].

Underspecified during access with later "homing-in", revisions are free within sentence [5, 6].

Immediately specified during access, revisions are costly.

We can sketch two reasons why underspecification may occur:

- **Utility**: Immediate commitment can result in costly reanalysis, and so should be avoided when possible.
- **Necessity**: Commitment to a narrow meaning for a polyseme cannot occur during a certain stage of incremental processing.

Evidence from a different reading task can help us tease these hypotheses apart.

The Maze task

In [7]'s Maze task, as implemented by [8], participants advance word-by-word by making decisions between the correct continuation of a sentence and a high-surprisal foil (see [9]).

- Choosing a foil terminates the trial.
- Success requires representation of structural and conceptual context.
- Response latency (RT) assumed to index lexical access, integration, and decision making.

Hypotheses & Predictions

Utility-based underspecification: Underspecification of polysemy is merely an effective option under typical comprehension strategies.

- Faced with a sufficiently different set of task pressures (e.g. the Maze's pressure for conceptual context), the processor may incidentally specify meanings for polysemes immediately.
- Polysemes will exhibit the same reanalysis costs as homonyms.

Necessary underspecification: Underspecification of polysemy is a requirement of lexical representation during processing.

- Despite the pressures of the task, the processor will still underspecify polysemes.
- Polysemes will, as in [1], exhibit less reanalysis cost than homonyms.

Methods & Results ($n = 48$)

Expanding on [1]'s design and stimuli, we collected word-by-word response latencies for 32 polysemy items and 32 homonymy items in the Maze, each crossing **Disambiguation Position** [Early, Late] × **Meaning** [M1 (dominant), M2 (non-dominant)].

Dominant meanings (M1) determined by new relative acceptability norming ($n = 32$), partially replicating [1].

Items Latin squared and randomized with 128 similar fillers, followed by a comprehension question.

Incomplete or incorrect trials excluded.

Dependent measure: sum of all log latencies in disambiguation region, residualized by position and characters.

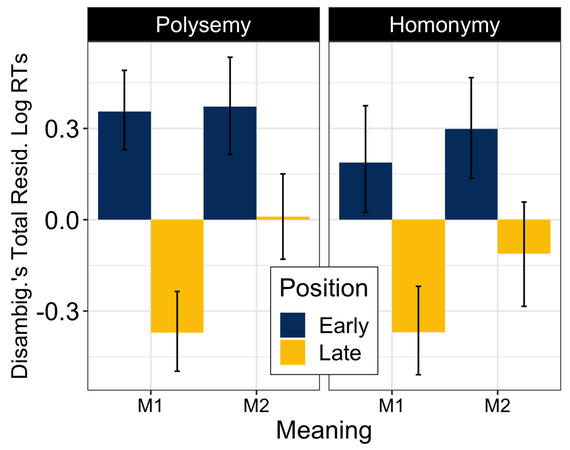
Polysemy	Meaning 1	Meaning 2	(Foil)
Early	Unfortunately, after it was soaked with rain the newspaper was destroyed.	Unfortunately, after it lost its advertising profits the newspaper was destroyed.	(x-x-x intend in job lips discover obtain kid conducted add extension.)
Late	Unfortunately, the newspaper was destroyed after it was soaked with rain.	Unfortunately, the newspaper was destroyed after it lost its advertising profits.	(x-x-x kid conducted add extension intend in job lips discover obtain.)
Homonymy			
Early	Reportedly, after it made his toast soggy the jam displeased Tom.	Reportedly, after it doubled his morning commute the jam displeased Tom.	(x-x-x come fit detail sir thinks begin kept ours indecision Need.)
Late	Reportedly, the jam displeased Tom after it made his toast soggy.	Reportedly, the jam displeased Tom after it doubled his morning commute.	(x-x-x kept ours indecision Need come fit detail sir thinks begin.)

Estimates and 95% CRIs from linear mixed models:

A main effect of Position such that Early disambiguators, which feature cataphora, feature slower RTs than Late disambiguators [10]. ☑ (-0.97,-0.50)

Utility hypothesis: A first-order interaction of Position × Meaning such that Late disambiguation to M2 is costly for all ambiguities. ☑ (0.05,0.66)

Necessity hypothesis: A second-order interaction of Position × Meaning × Target Type such that Late disambiguation to M2 is more costly for Homonymy than Polysemy. ☒ (-0.61, 0.29)



Discussion

The critical prediction of necessary underspecification is not met: we fail to replicate a difference in reanalysis costs between polysemy and homophony.

The presence of reanalysis costs for both ambiguities is **consistent with a utility-based account**: pressures of the Maze task lead the processor to commit fully, and ignore the option to delay.

Note that not all 'ambiguity' behaves alike: the ambiguity advantage for relative clause attachment is robust in the Maze [11]. Future use of the Maze may help draw clearer lines between effects attributable to processing strategy and the structure of the processor itself.

We could alternately interpret this as a straight-forward failure to replicate [1] in another modality. We aim to follow-up by verifying the effect's presence in self-paced reading and eyetracking using the same materials.

Conclusions

Differences in the online resolution of polysemy and homophony don't generalize from eyetracking to the Maze task.

We take the absence of this critical interaction to reflect the Maze task's pressure to maximize eager meaning commitment.

This leaves us with a theory of utility-based lexical underspecification rather than an account where it is necessary to delay commitment.

References

[1] Frazier, L., & Rayner, K. (1990). *JML, 29.*
[2] Frisson, S., & Pickering, M. J. (1999). *JEP: LMC, 25.*
[3] Frisson, S. (2009). *Lang Ling Compass, 3.*
[4] Klein, D. E., & Murphy, G. L. (2001). *JML, 45.*
[5] Foraker, S., & Murphy, G. L. (2012). *JML, 67.*
[6] Frisson, S., & Frazier, L. (2004). Poster at CUNY 17.
[7] Forster, K. I., Guerrera, C., & Elliot, L. (2009). *Behav Res Methods, 41.*
[8] Boyce, V., Futrell, R., & Levy, R. P. (2020). *JML, 111.*
[9] Sloggett, S., Van Handel, N., & Rysling, A. (2020a). Poster at CUNY 33.
[10] Filik, R., & Sanford, A. J. (2008). *Cognition, 107.*
[11] Sloggett, S., Van Handel, N., Sasaki, K., Duff, J., Rich, S., Orth, W., Anand, P., & Rysling, A. (2020b). Poster at CUNY 33.

Acknowledgements: We thank Jakub Dotlačil for filler materials, and Veronica Boyce, Shayne Sloggett, and Nick Van Handel for Maze resources.

Visit spellout.net/ibexexps/jmduff/FR-90-Maze-Sample/experiment.html for a Maze demonstration.