

Linguistic, but not visual, discourse context modulates syntactic parsing in the Maze

Veronica Boyce (Stanford University, vboyce@stanford.edu), Roger P. Levy (MIT)

How do comprehenders decide on parse structures when much linguistic input is temporally ambiguous? Language is usually intended as informative and communicative, so comprehenders can rely not only on language statistics and general knowledge, but also the *discourse context*, that is what entities are available in the visual or linguistic context. The use of contextual information has been most often studied with the visual world paradigm which can reveal the currently considered parse based on eye-movements^{1,2}. One open question is how rapidly different types of contextual information influence language processing, which can be measured using incremental processing measures to localize processing time effects. In the present work, we use the A-Maze task^{3,4} to explore the role of linguistic and visual discourse context in biasing syntactic attachment preferences. We describe two experiments testing NP versus VP preposition phrase attachment ambiguities in English (examples on reverse).

Materials, data, code for both experiments at github.com/vboyce/as-maze.

Experiment 1: Linguistic context We used the materials from Altmann & Steedman (1988)⁵ study 2 to see if the Maze task would reveal influences on attachment preferences based on the described context. In clause by clause SPR, A&S found an interaction where NP modifiers were preferred in the 2-referent condition (when there were two maps to distinguish between), but VP modifiers were preferred in the 1-referent condition (when there were a map and a manuscript, see reverse for full example text). We recruited 100 participants to each read 16 of the 32 items in an A-Maze paradigm (4 per condition). For the 88 participants who had greater than 90% overall selection accuracy, we found an expected interaction effect of 84 ms (95% CrI 17 – 148ms) on the disambiguating words (Table 1). Participants read words faster in the 2-referent NP and 1-referent VP conditions (Figure 1), indicating that preferred parse was influenced by whether the discourse context supported the use of a restrictive modifier.

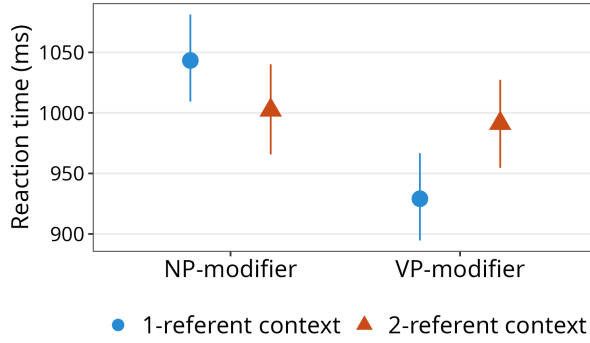
Experiment 2: Visual context Visually presented context affects online sentence parsing in visual world tasks^{1,2}. When two objects match the noun, the following PP is interpreted as a restrictive NP-modifier, judging by eye-movements. With only one object, the PP is initially interpreted as a VP-modifier goal leading to re-analysis when the actual goal PP occurs. More recent work has shown that visual context can facilitate language processing as measured by Maze RTs⁶. Here we ask if visual contexts that influence parsing in the visual world paradigm also affect Maze processing times. We used a Maze task where a context image was displayed for 5 seconds and then stayed visible as the Maze selections appeared below (example on reverse). We used a 2x2 design with 1- or 2- referent conditions in the visual context, and either ambiguous or unambiguous (“that’s” before first PP) sentences. We based our items on Weighall (2008)⁷; 200 participants each saw 4 critical items (1/condition) mixed with 4 fillers. 182 participants had accuracy above 90%. At the second preposition, there were effects of linguistic context (142ms longer RT [-2 – 258] in the ambiguous condition), but no effects of visual context (-27ms, [-99 – 45]) and no interaction (-34ms, [-146 – 86], Figure 2, Table 2).

Discussion: Many sources of information can constrain the meaning of incoming linguistic input, but it is unclear when different sources of information are incorporated. Additionally, task demands may affect what information sources seem most relevant and how much information can be incorporated quickly. Here, we demonstrate that, for a prepositional phrase attachment ambiguity, linguistically presented discourse context does affect processing times during the Maze task. Co-presented visual context did not affect Maze reading times, in contrast to its effects on eye movements in visual world paradigms. This difference could be due to high cognitive load from Maze and lack of task demands to closely inspect the image, or the two tasks could be querying different stages in linguistic processing and interpretation. Future work could explore whether visual discourse context influences processing times in other experimental paradigms.

Experiment 1: Linguistic context

Example item: (*context sentence*) A historian was working in the British Museum holding a magnifying glass. He'd sat down to study a map. (*1 or 2 referents introduced*) On his desk there was **a map** which had an appalling tear and **a [manuscript | map]** which seemed in perfect condition. (*a PP that resolves to NP or VP attachment, disambiguation underlined*) The historian had to study **the map with the [appalling tear | magnifying glass]** so as to value it.

Figure 1: Expt 1 Results



Means and bootstrapped 95% CIs for each of the 4 conditions. Data from 88 participants who each completed 16 items. There is an interaction between the context and what the PP modifies.

Table 1: Expt 1 model coefficients

Term	Est.	95% CrI
Intercept	995 ms	[943 – 1046]
Type (VP)	-73 ms	[-134 – -11]
Context (2-referent)	22 ms	[-13 – 59]
Type x Context	84 ms	[17 – 148]

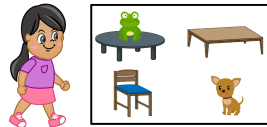
Model: $RT \sim type \times context + (type \times context / item) + (type \times context / subject)$ Fit to reading times on the disambiguating words for correct responses only. Type and context were sum coded (VP, 2-referent coded as +.5; NP, 1-referent as -.5) and weakly regularizing priors were used.

Experiment 2: Visual context

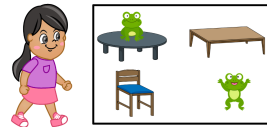
Example item:

Visual context:

(seen for 5 seconds before starting sentence)



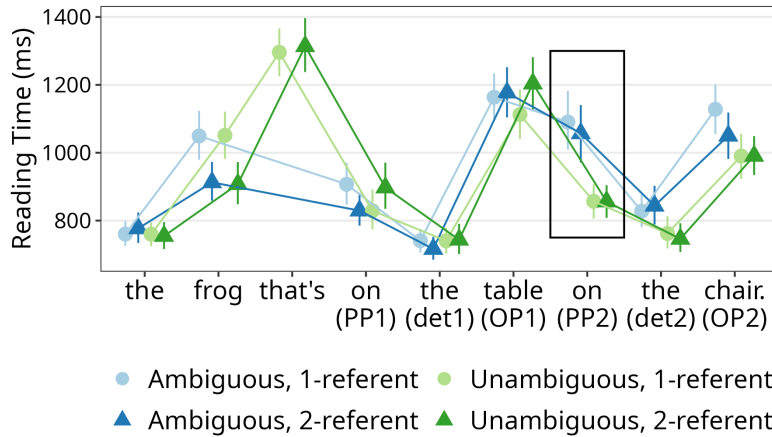
1-referent (frog and dog)



2-referent (2 frogs)

Sentence (*ambiguous v unambiguous*): The girl will put the frog [**that's**] on the table on the chair.

Figure 2: Expt 2 Results



Estimate and bootstrapped 95% CI for each of the 4 conditions. The second preposition (boxed) is the critical position. Data from 182 participants who each saw 4 critical items. There is a main effect of sentence type, where ambiguous sentences have a longer RTs on the second PP, but no interaction with visual context.

Table 2: Expt 2 model coefficients

Term	Est.	95% CrI
Intercept	931 ms	[811 – 1041]
Type (Ambiguous)	142 ms	[-2 – 258]
Context (2-referent)	-27 ms	[-99 – 45]
Type x Context	-34 ms	[-146 – 86]

Model: $RT \sim type \times context + (type \times context / item) + (type \times context / subject)$ Fit to reading times on the second preposition for correct responses only. Type and context were sum coded (Ambiguous, 2-referent coded as +.5; unambiguous, 1-referent as -.5) and weakly regularizing priors were used.

References: • 1. Tanenhaus et al. Science, 1995. • 2. Trueswell et al. Cognition, 1999. • 3. Boyce, Futrell, Levy. JML, 2020. • 4. Boyce, Levy. Glossa Psycholinguistics, 2023. • 5. Altmann, Steedman. Cognition, 1988. • 6. Pushpita & Levy. CoNLL, 2024. • 7. Weighall. J Exp Child Psych, 2008.