Reanalysis in discourse comprehension: Evidence from reading times

Introduction: This paper explores 'discourse reanalysis', where in a sequence of discourse units $S_a ... S_x$, an initial interpretation of S_a must be revised given S_x. E.g., at S_2 of (1), the narrative seems to describe a series of parallel events, where *him* picks out George. But given S_3, *him* must instead pick out Bill, and S_1 and S_2 would now seem to describe an action and a reaction.

(1) [Bill pinched George]$_{S_1}$ [and Ava elbowed him.]$_{S_2}$ [Bill didn't elbow her in return,]$_{S_3}$...

As noted by [9], such a discourse poses problems for logics that don't permit selective non-monotonic update of coreference—e.g., in DRT [12], once a pronoun *him* has been represented as coreferent with its antecedent, the downstream semantics cannot distinguish propositions involving the uncertain referent of *him* from those involving the antecedent. This observation motivated [9] to develop a variant of DRT (PCDRT) with selective non-monotonic update of coreference (see also [1, 10, 11, 14] for other non-monotonic approaches).

One objection to theories featuring a selective non-monotonic update is that a simpler alternative looms: a fully-underspecified approach (see [4] for discussion), which could maintain all grammatical discourse interpretations, without any selection or ranking to later be revised. To see which approach best models the representations used in actual incremental comprehension, we turn to psycholinguistic evidence. The revision expected by the non-monotonic approach should involve momentary difficulty akin to 'garden path' effects in syntactic processing [5, 6]. To date, psycholinguistic studies have generally failed to find such reanalysis costs for discourse, leaving it possible that readers may underspecify interpretations of e.g. ambiguous pronouns [20], tense [2, 18], and coherence relations [3, 16]. The goal of the present work is to show, for the first time, that there is evidence for costly reanalysis related to ambiguities of coreference and coherence in discourses like (1). We take this to solidify the empirical basis for non-monotonic update of discourse-pragmatic content, with important consequence for formal theories.

Experiment 1: Following [13, 19, 21], we set up 60 narratives featuring coordinated clauses with parallel syntactic structure (S_1, S_2), where S_2 contained a potentially-ambiguous object pronoun. The interpretation of such pronouns is biased towards an object antecedent due to parallelism [19], but this can be reversed by a preceding causal connective [21], e.g. *so*, or by using a second verb biased more towards the causal relation Result than Parallel [13]. We elicited preferred interpretations of the object pronoun, and judgments of certainty, across six key conditions (Table 1) to verify the contributions of the verb and connective, and further examine how preferred interpretations changed when we added an S_3 where the presupposition of *in return* disambiguated the previous pronoun.

We tested 30 native English speakers recruited on Prolific (Figure 1). We analyzed responses as an ordinal variable combining preference and certainty data, using Bayesian ordinal mixed-effects regression. Models confirmed a strong preference for an object antecedent in the ambiguous, Parallel condition (80%, with moderate certainty ratings), which was sharply reversed by causal-biased verbs, $\beta_{.95} = (-1.92, -1.24)$, by the causal connective *so*, $\beta_{.95} = (-2.05, -1.31)$, and, most strongly, by an S_3 disambiguating to a subject antecedent, $\beta_{.95} = (-3.34, -2.31)$.

These results replicate a strong parallelism-based object antecedent bias in such narratives, and further confirm that preferred interpretations of ambiguous pronouns can be reversed by a subsequent sentence. Note that in this design, and in Exp2, data cannot distinguish between a shift in coreference and coherence together, or a shift in coreference alone; see discussion below.

Experiment 2: From Exp1, we identified 32 narratives with the strongest parallelism bias in S_2 and the largest reversal given S_3, to investigate in a chunked, non-cumulative self-paced-reading experiment testing for costly reanalysis. We compared these 32 (mixed with 60 fillers) across four conditions (Table 2). In Ambiguous conditions, the critical object pronoun is expected to prefer an initial object antecedent due to establishing Parallel, but it is followed by an S_3 which requires a subject antecedent. In all other conditions, S_3 doesn't motivate any change in interpretation, because S_2 already makes a subject antecedent unambiguous, or because S_2

lacks a pronoun. If comprehenders indeed commit to a firm interpretation of S_2 that they must then reanalyze, we should see particularly slow reading of S_3 in Ambiguous conditions.

Reading times (Figure 2) from 96 participants (recruited as in Exp1) were analyzed using Bayesian log-normal mixed-effects regressions. We find that the final sentences were generally read slightly slower in the Ambiguous condition (red) than the other conditions, with our models estimating credible marginal slowdowns of about 20ms at *in return*, $P(\Delta > 0) = 0.97$, and *she simply*, $P(\Delta > 0) = 0.98$. This result suggests that readers experienced costly reanalysis, revising an initial interpretation that yielded an object antecedent for *her*, to a new interpretation delivering the newly-mandated subject antecedent. Evidence for costly revision is strongest for the first trials of the experiment, and disappears by the last few, resembling adaptation effects well-known in work on syntactic garden paths [17]. This cost may include revision of an initial parallel relation (Parallel) to a causal relation (Result); theories and intuitions suggest this should also occur, but its costs cannot be untangled from pronominal revision.

Reading times on the preceding region offer further insight. Besides costs for a new discourse referent (purple), we also see here an inversion of the later pattern; conditions with unambiguous subject antecedents (green, blue) were slower than Ambiguous (red), $P(\Delta > 0) = 0.93$. This isn't a simple effect of disambiguation, cf. [8], as unambiguous object antecedent conditions in our filler items are not similarly slow. We take this to indicate that participants begin reading this region already expecting Parallel and thus parallelism in coreference, and their comprehension is facilitated when these expectations are met, as [16] finds when Explanation is expected.

Discussion: The reading data from Exp2 provides evidence for a slowdown just when comprehenders encounter input that encourages a previously despised set of coreference and coherence choices. We take this to diagnose costly reanalysis of discourse interpretation, consistent with the idea that, at least in some cases, comprehenders make and revise firm pragmatic representations during incremental discourse comprehension, and do not somehow avoid revision through prolonged underspecification. This joins a few other cases showing reanalysis costs for pronouns in simpler narratives [7, 15]. Other work [2, 20] has perhaps failed to observe such costs because they have examined cases with weaker biases, where comprehenders may underspecify [3].

The consequences of our results for theories of semantic representation depend on one's theoretical taste. One could attempt to reconcile these reanalysis costs with those seen for syntactic or lexical ambiguity, e.g. arguing that they involve wholesale reconstruction of an alternative analysis permitted by the grammar, as carried out by a separate meta-system aware of the constraints from later input (a 'semantic parser'). This would require no particular flexibility for discourse-pragmatic meaning in the semantic representations themselves. A theory which desires a single, cognitively-adequate logic, on the other hand, will indeed need to handle non-monotonic update in one of the ways discussed by [1, 9–11, 14]. The latter approach might have one empirical advantage: if our 20ms estimate here is generalizable, this reanalysis seems to be somewhat less difficult than syntactic garden paths observed in self-paced reading, e.g., 50ms in [5]. This would be easier to understand if discourse reanalysis is unique in that it does not require altering the grammatical representation of the input—and this would require grammatical representations which separate discourse-pragmatic content into a non-monotonic logic.

Next steps: If costly reanalysis is a general matter of discourse comprehension, we predict similar effects when coherence drives resolution of e.g. temporal ordering. To this end, our next experiments will test cases like (2), where S_1 and S_2 first appear to relate via Explanation, such that the event described by S_2 would be understood to precede the event described by S_1. S_3 would then force reanalysis to Result, ensuring the opposite event ordering. Such a reanalysis would be from a subordinating relation to a coordinating one, in the sense of [1], whereas our present study featured only coordinating relations. If coherence revision is a source of difficulty, as we hypothesize, this larger structural shift may drive larger reanalysis costs.

(2) [Bob sued Liz.]$_{S_1}$ [She defaced his shop,]$_{S_2}$ [because his lawsuit was cruel.]$_{S_3}$

Preregistrations, materials, and analysis for both studies are available in our anonymous OSF repository.

The kids were misbehaving at a fancy dinner.		
Pronoun	**Context**	**Text**
Ambiguous	Parallel-biased V	Mia hit Winona with a pea and Ian kicked her under the table.
	Result-biased V	Mia hit Winona with a pea and Ian scolded her.
	Result (*so*)	Mia hit Winona with a pea so Ian kicked her under the table.
Disambig (S)	Parallel-biased V	Mia hit Winona with a pea and Ian kicked her under the table. Mia didn't kick him in return, she simply...
Disambig (O)	Parallel-biased V	Mia hit Winona with a pea and Ian kicked her under the table. Winona didn't bother the other kids in return, she simply...
Unambig (S/O)	Parallel-biased V	Mia hit Harrison with a pea and Ian kicked (her/him) under the table.
Is it more likely that Ian kicked Mia or Winona? (Mia, Winona) *How likely is the other option?* (1-4)		

Table 1: Stimuli from Experiment 1.

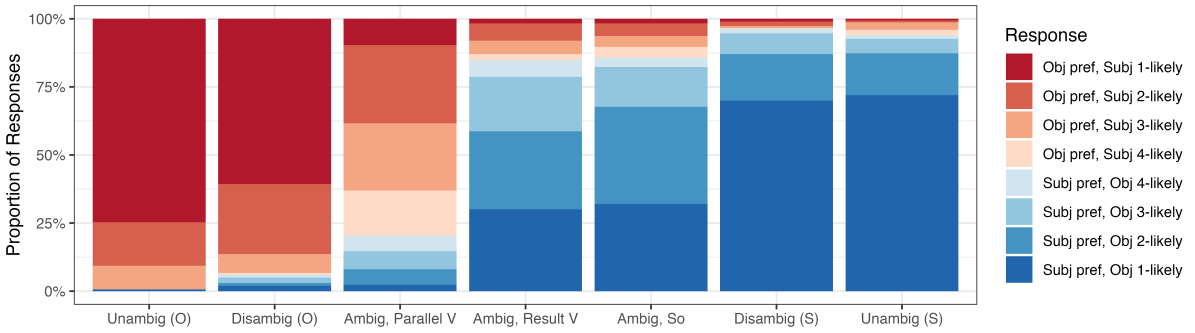

Figure 1: Results from Experiment 1.

The kids/were misbehaving/at a fancy dinner.	
Ambig.	Mia hit Winona/with a pea/and/**Ian kicked her**/under the table.
Unambig.	Mia hit Harrison/with a pea/and/**Ian kicked her**/under the table.
so + Unambig.	Mia hit Harrison/with a pea/so/**Ian kicked her**/under the table.
Name	Winona hit Harrison/with a pea/and/**Ian kicked Mia**/under the table.
Mia didn't kick him/**in return**,/**she simply**/went to get more peas.	

Table 2: Stimuli from Experiment 2. "/" indicates a chunk boundary. Critical regions in bold.

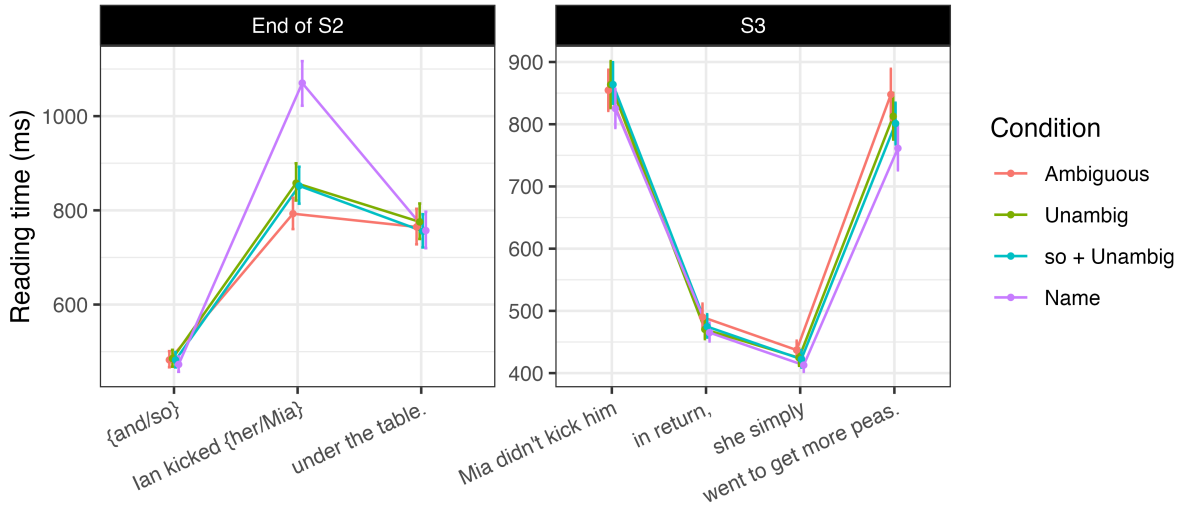

Figure 2: Results from Experiment 2.

References

[1] Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge.

[2] Dickey, M. W. (2001). *The processing of tense*. Kluwer.

[3] Duff, J. (2023). *On the timing of decisions about meaning during incremental comprehension*. Dissertation, UC Santa Cruz.

[4] Egg, M. (2010). Semantic underspecification. *Language and Linguistics Compass*, *4*(3).

[5] Ferreira, F., & Henderson, J. M. (1990). Use of verb expectation in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4).

[6] Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*(2).

[7] Gordon, P. C., & Scearce, K. A. (1995). Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory & Cognition*, *23*(3).

[8] Grant, M., Sloggett, S., & Dillon, B. (2020). Processing ambiguities in attachment and pronominal reference. *Glossa*, *5*(1).

[9] Haug, D. T. T. (2014). Partial dynamic semantics for anaphora: Compositionality without syntactic coindexation. *Journal of Semantics*, *31*.

[10] Jaspars, J., & Kameyama, M. (1998). Discourse preferences in dynamic logic. In A. Aliseda, R. van Glabbeek, & D. Westertahl (Eds.), *Computing natural language*. CSLI.

[11] Kameyama, M. (1996). Indefeasible semantics and defeasible pragmatics. In M. Kanazawa, C. Piñón, & H. de Swart (Eds.), *Quantifiers, deduction, and context*. CSLI.

[12] Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to model theoretic semantics of natural language, formal logic and discourse representation theory*. Kluwer.

[13] Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25*(1).

[14] Lascarides, A., & Asher, N. (1993). Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy*, *16*.

[15] Mak, W. M., & Sanders, T. J. M. (2010). Incremental discourse processing: How coherence relations influence the resolution of pronouns. In M. Everaert, T. Lentz, H. De Mulder, O. Nilsen, & A. Zondervan (Eds.), *The linguistics enterprise: From knowledge of language to knowledge in linguistics*. John Benjamins.

[16] Mak, W. M., & Sanders, T. J. M. (2013). The role of causality in discourse processing: Effects of expectation and coherence relations. *Language and Cognitive Processes*, *28*(9).

[17] Prasad, G., & Linzen, T. (2021). Rapid syntactic adaptation in self-paced reading: Detectable, but only with many participants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(7).

[18] Sasaki, K. (2021). *Components of coherence*. Dissertation, UC Santa Cruz.

[19] Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, *23*(3).

[20] Stewart, A. J., Holler, J., & Kidd, E. (2007). Shallow processing of ambiguous pronouns: Evidence for delay. *Quarterly Journal of Experimental Psychology*, *60*(12).

[21] Wolf, F., Gibson, E., & Desmet, T. (2004). Discourse coherence and pronoun resolution. *Language and Cognitive Processes*, *19*(6).