

Reading measures as a window into QUD accommodation: Evidence from corrections

The notion of Question-Under-Discussion (QUD) has proven useful in accounting for various phenomena [1-8], suggesting comprehenders track one or more potentially hierarchically organized discourse goals—*strategies of inquiry*—through which overarching superquestions are resolved by a series of subquestions [5-8] (Fig1). Since QUDs are often left implicit, they must be inferred based on linguistic devices conventionally tied to a specific QUD, like focus marking. The accommodation of these complex structures may pose a challenge for comprehenders: QUDs can shift rapidly across utterances, often requiring novel QUDs to be integrated before previous ones are resolved, as in (1b) vs. (1a). Building on work studying focus in reading [9-12], we investigate the cost associated with the accommodation of novel QUDs, and how it is modulated by the larger discourse structure. While prior work has examined how QUDs established earlier in the context affect processing of subsequent material [e.g., 18-20], this study, to our knowledge, offers the first evidence on the online inference of QUD structure itself.

We specifically tested reading of corrections [13-14], a conversational move triggered by the presence of focus which either corrects a prior answer to the same QUD (1a) or shifts the discourse to a novel one (1b). If shifting towards a novel QUD incurs additional processing costs (H_1 : NOVELTY COST), RTs on the focus should be longer in (1b) than (1a). We further compared corrections to moves where a newly inferred QUD is parallel to an earlier one, as when a contrastive topic (CT) is followed by a contrastive focus (1c). CTs are standardly analyzed as signaling the relevance of a parallel question under the same superordinate QUD [6-8]. Unlike in (1c), the QUD in (1b) cannot be taken to maintain the existing strategy, because it shifts towards asking about each individual who invited them, instead of who everyone invited (Fig1). Thus, if maintaining a strategy aids the accommodation of novel QUDs (H_2 : STRUCTURE BENEFIT), RTs should be shorter in (1c) than in (1b). **Methods.** H_1 and H_2 were tested using the Maze task [15-16], in which each item ($n=48$) comprised a three-turn dialogue, preceded by a short description of background assumptions, as in (2). The first two utterances in each dialogue manipulated the Type of focus in the target sentence (CONTRAST vs. CORRECT), and QUD Novelty (SAME vs. NOVEL), holding the final target sentence constant across conditions (except for the inclusion of the particle *too* to ensure felicity in (2c)). In the SAME conditions, only the final noun (NP2, *cookbooks*) was new, maintaining the previous QUD; in the NOVEL conditions both NP2 and subject (*Lucy*) were new, triggering a shift towards a novel QUD upon reading the subject. Subjects in the NOVEL CONTRAST condition were CTs, while they were QUD-shifting corrections in the NOVEL CORRECT condition. All contexts were presented normally, with only targets presented word-by-word (Fig2). Log-RTs on subject and NP2 from 60 participants were analyzed using Bayesian mixed effects models ([17], Tab1). **Results.** Models revealed a reliable interaction on the subject, indicating faster RTs on CTs than on QUD-shifting foci. Results also revealed a main effect of Type and an interaction effect on NP2, indicating longer RTs in the NOVEL CORRECT than both the NOVEL CONTRAST ($\beta=.11(.02)$, 95%CrI=[.07,.15]) and SAME CORRECT conditions ($\beta=.07(.02)$, 95%CrI=[.03,.11]). The difference among the SAME conditions was not reliable, however, potentially because comprehenders anticipated the additive particle in the SAME CONTRAST condition, which only appeared after NP2.

In sum, results suggest a NOVELTY COST, with greater slowdowns for QUD-shifting than same-QUD foci, confirming H_1 . Results also indicated a STRUCTURE BENEFIT, with entirely novel QUDs being costlier than novel QUDs that fit within an established strategy of inquiry, as in H_2 . This suggests that accommodation cost is larger the more the inferred QUD deviated from established structure. Thus, the possibility of inferring such strategies does not pose a burden, but instead facilitates the accommodation of specific QUDs. Given that foci were new in all conditions, these effects can only be due to the relationship between incoming information and context. Overall, this pattern therefore suggests that comprehenders do not just incrementally update immediate QUDs, but also track how they fit into the overall discourse structure.

- (1) a. **A:** Who did Lily invite?
B: Lily invited PETER_F
C: No, Lily invited **BILL_F** *same-QUD correction*
↳ QUD: Who did Lily invite?
- b. **A:** Who did Sam invite?
B: Sam invited PETER_F
C: No, **LILY_F** invited Bill *novel-QUD-shifting correction*
↳ QUD: Who invited Bill?
- c. **A:** Who did Sam invite?
B: Sam invited PETER_F
C: And **LILY_{CT}** invited **BILL_F** *parallel-QUD-shifting move*
↳ QUD: Who did Lily invite?

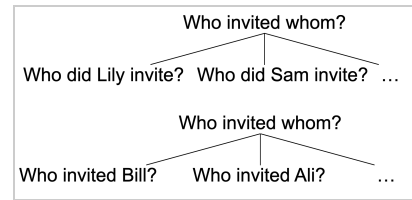


Figure 1. Two strategies of inquiry for one superquestion

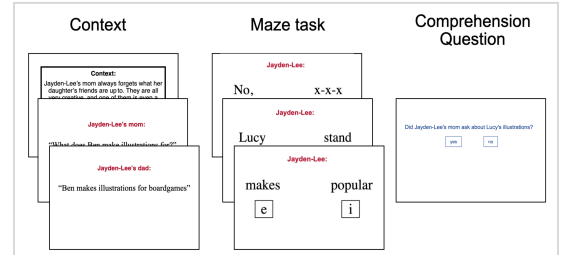


Figure 2. Trial structure

- (2) **Context:** Jayden-Lee's mom always forgets what her daughter's friends are up to. They are all very creative, and one of them is even a professional illustrator.
- a. **A:** What does Lucy make illustrations for?
B: Lucy makes illustrations for [board games]_F.
C (target): No, Lucy makes illustrations for [**cookbooks**]_F, I think. SAME CORRECT
- b. **A:** What does Ben make illustrations for?
B: Ben makes illustrations for [board games]_F.
C (target): No, [**Lucy**]_F makes illustrations for [**cookbooks**]_F, I think. NOVEL CORRECT
- c. **A:** What does Lucy make illustrations for?
B: Lucy makes illustrations for [board games]_F.
C (target): And Lucy makes illustrations for [**cookbooks**]_F, too, I think. SAME CONTRAST
- d. **A:** What does Ben make illustrations for?
B: Ben makes illustrations for [board games]_F.
C (target): And [**Lucy**]_{CT} makes illustrations for [**cookbooks**]_F, I think. NOVEL CONTRAST

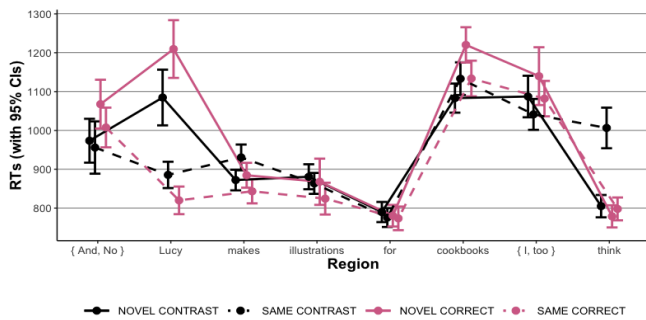


Figure 3. RTs per condition across regions

	Subject (<i>Lily</i>)		NP2 (<i>cookbooks</i>)	
	β	(error) 95% Cr.I.	β	(error) 95% Cr.I.
Intcpt	6.78 (0.04)	[6.69, 6.86]	6.96 (0.04)	[6.88, 7.04]
TYPE	0.00 (0.02)	[-.03, 0.03]	0.05 (0.01)	[0.02, 0.08]
NOVELTY	0.25 (0.02)	[0.21, 0.29]	0.01 (0.02)	[-.02, 0.05]
TY x NOV	0.18 (0.04)	[0.10, 0.25]	0.11 (0.03)	[0.05, 0.16]

Table 1: Posterior estimates per region per condition (logRTs)

References. [1] Zondervan (2008) *SALT* [2] Onea (2016) *Brill*. [3] Tonhauser et al. (2013) *Language*. [4] Simons et al. (2010) *SALT*. [5] Beaver & Clark (2008) Oxford: Blackwell. [6] Roberts (2012) *Sem. & Prag*. [7] Buring (2003) *Ling. & Phil*. [8] Constant (2014). UMass thesis. [9] Birch & Rayner (1997) *Mem. & Cogn*. [10] Benatar & Clifton (2013) *JML*. [11] Lowder & Gordon (2015) *Psych. Bulletin & Review*. [12] Hoeks et al., (2023) *JML*. [13] Jasinskaja (2012) *Lingua*. [14] Jasinskaja (2013) *Lingua*. [15] Forster et al. (2009) *Behav. Res. Methods*. [16] Boyce et al. (2020) *JML*. [17] Bürkner (2017) *J. Stat. Soft.* [18] Clifton & Frazier (2012) *Cogn. Psych*. [19] Kim (2012) Rochester thesis. [20] Göbel (2020) UMass thesis