Published in final edited form as:

J Mem Lang. 2016 January 1; 86: 20–34. doi:10.1016/j.jml.2015.10.004.

Accommodation to an Unlikely Episodic State

Charles Clifton Jr.^a and Lyn Frazier^b

^a Department of Psychological and Brain Sciences, University of Massachusetts Amherst, Amherst, MA 01003 USA

^b Department of Linguistics, University of Massachusetts Amherst, Amherst, MA 01003 USA; lyn@linguist.umass.edu

Abstract

Mini-discourses like (ia) seem slightly odd compared to their counterparts containing a conjunction (ib).

(i) a. Speaker A: John or Bill left.

Speaker B: Sam did too.

b. Speaker A: John and Bill left. Speaker B: Sam did too.

One possibility is that *or* in Speaker A's utterance in (ia) raises the potential Question Under Discussion (QUD) whether it was John or Bill who left and Speaker B's reply fails to address this QUD. A different possibility is that the epistemic state of the speaker of (ia) is somewhat unlikely or uneven: the speaker knows that someone left, and that it was John or Bill, but doesn't know which one. The results of four acceptability judgment studies confirmed that (ia) is less good or coherent than (ib) (Experiment 1), but not due to failure to address the QUD implicitly introduced by the disjunction because the penalty for disjunction persisted even in the presence of a different overt QUD (Experiment 2) and even when there was no reply to Speaker A (Experiment 3). The hypothesis that accommodating an unusual epistemic state might underlie the lower acceptability of disjunction was supported by the fact that the disjunction penalty is larger in past tense discourses than in future discourses, where partial knowledge of events is the norm (Experiment 4). The results of an eye tracking study revealed a penalty for disjunction relative to conjunction that was significantly smaller when a lead in (*I wonder if it was...*) explicitly introduced the disjunction. This interaction (connective X lead in) appeared in early measures on the disjunctive phrase itself, suggesting that the input is related to an inferred epistemic state of the speaker in a rapid and ongoing fashion.

It has long been established that a sentence seems most natural, and is most easily comprehended, when it meshes easily with the current shared model of the discourse in

Corresponding author: Charles Clifton Jr., Department of Psychological and Brain Sciences, University of Massachusetts Amherst, Amherst, MA 01003 USA (cec@psych.umass.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

which it occurs (Bock & Mazzella, 1983; Clark & Haviland, 1977; Haviland & Clark, 1974; Schwarz, 2015a). Sometimes a sentence requires that the discourse model must be adjusted to make sense of the sentence. In such a case, 'accommodation' is said to take place. Presupposition accommodation is one much-studied instance of accommodation where comprehenders add relevant information to the common ground even though it is only presupposed and not asserted. For instance, hearing *John's son goes to UMass* might lead a comprehender to add the previously unknown information that John has a son of college age, or at least accept that the speaker is speaking as if John has a son (Beaver & Zeevat, 2007; Lewis, 1978; Schwarz, 2015b).

Beyond presupposition accommodation, there are other instances of accommodation where comprehension depends on developing a context shared between speaker and hearer that extends beyond explicitly conveyed information. In this paper, we consider two such instances. The first assumes that the goal of a sentence or discourse is to work toward an answer of a question, explicit or implicit, potentially as broad as "what is the state of things?" (Roberts, 1996/2012). A sentence that addresses a question that differs from the question suggested by the preceding discourse may be difficult to integrate with the discourse; accommodation to a new set of speaker's goals may be required. We discuss this "Question Under Discussion" (QUD) suggestion below. The second kind of accommodation comes about when a sentence seems to clash with the state of knowledge that the speaker or writer is likely to be in. We term this suggestion the "epistemic state" hypothesis, and discuss it further below. Following a brief description of one line of psycholinguistic research on the processing costs of accommodation, we describe these two kinds of accommodation in more detail, reviewing evidence for their effects. We then turn to a novel phenomenon, difficulty of processing disjunctive vs. conjunctive noun phrases, and explore in a series of experiment whether one or both of these kinds of accommodation is responsible for the phenomenon.

Perhaps the most-studied (Clifton, 2013; Frazier, 2006; Garrod & Sanford, 1982; Haviland & Clark, 1974; Singh, Fedorenko, Mahowald, & Gibson, 2015 in press) instance of accommodation to the epistemic state of the speaker involves comprehension of sentences with definite vs indefinite articles. Use of *the* presupposes that the speaker has in mind a specific or a familiar referent; use of *a* presupposes that this is not the case, and that there are possibly multiple possible referents (see Hawkins, 1978, for discussion). When the discourse context does not support the familiarity presupposition, a reader or listener must accommodate it, at some cost in comprehension ease. Clifton (2013) showed that there was a processing cost when the presupposition of either the definite or the indefinite was made unlikely by the content of a sentence (e.g., reading was slowed in *In the kitchen, Jason checked out a stove*, compared to *...the stove*). Singh, Fedorenko, Mahowald, & Gibson (2015, in press) went further and showed that accommodating a presupposition in a context that made it implausible was even more costly than reading an assertion of the same content.

Despite such demonstrations of the cognitive cost of having to accommodate a difficult presupposition, a surprising fact about language is that the acceptability of sentences or mini-discourses is sometimes improved, not degraded, when accommodation is necessary. We suggest that it is more felicitous to say *John's wife is a chemist* than *John has a wife*.

She's a chemist (a judgment that has received some support in informal discussions). This is presumably true only if it is easy for the comprehender to imagine the situation or epistemic state of the speaker that is implied, and it suggests that such well-supported accommodation facilitates communication (see Piantadosi, Tily, & Gibson, 2012, for related discussion).

The QUD

We now turn to a discussion of the first type of accommodation we introduced above, accommodation to the apparent communicative goals of a speaker or writer, or alternatively, accommodation of a changed QUD. The basic idea is that discourse is structured as a series of explicit or implicit questions followed by answers or comments on the current question. The notion of Question Under Discussion (QUD) has played an important role in a variety of areas in linguistics (Beaver & Clark, 2008; Ginzburg, 2012; Roberts, 2012/1996. For instance, influential proposals cash out the notion 'discourse topic' in terms of the current QUD (van Kuppevelt, 1996).

Formal semantic theories of questions treat a question as the set of its possible or true answers (Hamblin, 1973; Karttunen, 1977). This semantic analysis of questions makes it clear how questions can play a powerful role in organizing discourse. A coherent discourse progresses by supporting the choice of one or more propositions from the set of alternatives that are introduced by an implicit or explicit question (the QUD). A focused constituent can be viewed as selecting from among a set of alternatives (Rooth, 1985, 1992) and thus answering a QUD. Sometimes the QUD can be identified only after-the-fact, but sometimes a discourse introduces a QUD that can guide the analysis of upcoming material.

A variety of experimental demonstrations suggest that processing of a discourse is facilitated when material addresses a likely QUD, and impaired when material requires that the comprehender replace the likely QUD (an instance of accommodation, in present terms). For instance, Grant, Clifton, and Frazier (2012) showed that elliptical sentences which they viewed as ungrammatical, such as *The information was released but Gorbachev didn't*, were judged to be less unacceptable if they contained a modal (e.g., *The information needed to be released, but Gorbachev didn't*). They argued that this modal raised a question like "Was it or wasn't it?" which the reader could take as the QUD, which would facilitate comprehension of the following elliptical phrase. (See Clifton & Frazier, 2012, Experiments 2 and 3, for similar evidence using eyetracking while reading.) Tian, Breheny, and Ferguson (2010) explicitly used the notion of QUD to account for a very different phenomenon. They measured the time to decide whether or not a pictured object had been mentioned in a previously read sentence, and found that decision times were faster when the state of the pictured object was congruent with the state implied by the apparent QUD of the sentence. For instance, subjects were faster to say 'yes' to a picture of a pile of uncooked spaghetti than to a pile of cooked spaghetti after hearing *It was Jane who didn't cook the spaghetti*, a sentence that presumably suggested the implicit QUD "Who didn't cook the spaghetti?"

Zondervan, Meroni, and Gualmini (2008) studied the effects of an explicit QUD. They showed that the scope ambiguity of *All the pizzas were not delivered* was resolved differently following the question *Were all the pizzas delivered?* than following the question

Were some pizzas delivered? (when shown that two of four pizzas were delivered, subjects typically said "true" to the *All the pizzas* sentence following the first question, but "false" following the second). Explicit context questions similarly affected whether listeners made a scalar implicature to *I think some hotdogs were delivered* when judging the truth of that sentence in a situation where all the hotdogs were delivered. While not explicitly invoking the QUD notion, Rohde and Horton (2014) used a sort of visual world 'game' to explore how a discourse introduces implicit questions that can be answered by later material. The discourse context consisted of a sentence that implicitly introduced the question "why" (e.g., *Beryl delighted John because she applauded him*) or one that implicitly introduced the question "what next." Subjects who were successfully trained to identify discourses as containing causal or continuation relationships showed (by where they looked on the game's visual display) that they anticipated a causal continuation to the discourses that implicitly asked "why" and a consequence continuation to those that implicitly asked "what next."

Finally, Clifton and Frazier (2012) presented eye tracking evidence showing that readers are slow in reading a region of a sentence where a tempting explicit QUD predicts one analysis of a sentence but then the sentence proves to be incompatible with that tempting QUD. Their subjects read a context question that contained a verb that is biased (in corpus counts and completion norms) toward having a NP complement (direct object), such as *write*, or one that is biased toward having a S complement, such as *say* (e.g., *What did the CIA director say?*) They then read a sentence that was disambiguated toward a direct object vs. a S complement answer (He confirmed the rumor once it had spread widely/He confirmed the rumor should have been stopped sooner). Reading was disrupted (as reflected primarily in regressions from the disambiguating region) when the syntactic resolution clashed with the bias of the verb in the context question, presumably because readers had to accommodate an unexpected QUD.

Epistemic state

The other kind of accommodation introduced earlier involves changes in the knowledge that can be assumed to be available to a speaker or writer. We refer to this as the 'epistemic state' of the speaker or writer. Readers and listeners' interpretations of a sentence clearly can depend on the knowledge state they attribute to its author. A response to a request for directions on how to reach a destination is more likely to be followed if it is delivered in a confident manner than if it is delivered in a hesitant fashion. Grodner and Sedivy (2015, in press) have demonstrated a similar effect by manipulating the apparent competence of a speaker. Listeners in a visual world experiment, instructed to (e.g.) *Move the tall cup*, showed more sensitivity to contextual contrast when the speaker was purportedly competent than when the speaker was described as being cognitively impaired and made apparent errors. Similarly, Bergen and Grodner (2012) showed that readers were more likely to generate scalar implicatures (e.g., *some* implicates *not all*) when the discourse containing the scalar term suggested that its author had relevant knowledge (e.g., would have known where *all* was appropriate). Breheny, Ferguson, and Katsos (2013) used a visual world technique to show that listeners were more likely to make ad-hoc 'and nothing else' inferences about a previously-seen action that was being described by an individual if they knew that that individual has seen all of the action rather than only part of it. In this paper, we explore the

possibility that a sentence that requires an unusual or uneven epistemic state to be attributed to a speaker or writer makes that sentence difficult to process.

The current experiments

In the present paper, we introduce a phenomenon that, we propose, could potentially be accounted for by either (or both) of the two kinds of accommodation we have introduced. The phenomenon involves the processing of conjunctive vs. disjunctive noun phrases. We observed (as will be described in detail in Experiment 1) that dialogs like those in (1) are judged to be less natural and less acceptable when the connective is *or* (as in (1a)) than when it is *and* (1b).

(1) a. Speaker A: John or Bill left.

Speaker B: Sam did too.

b. Speaker A: John and Bill left.

Speaker B: Sam did too.

The relative unnaturalness of the discourses with the disjunction could involve the appropriateness of Speaker B's response to the apparent QUD. In (1a), we assume that a listener or reader will, following Speaker A's contribution, accommodate a QUD along the lines of "Which of John or Bill left?" Speaker B's response does not address this QUD, and hence may seem infelicitous. In contrast, in (1b) the natural QUD would be something like "Who left?" and Speaker B's response does address this question.

Alternatively, or additionally, the relative unnaturalness of (1a) could reflect the plausibility of the epistemic state of the speaker who uses one of the connectives. In order to say *X or Y verbed*, a speaker must be in a peculiar knowledge state. The speaker knows that some people engaged in the activity, and knows who they might be, but does not know who they are. The difficulty of accommodating such an epistemic state of the speaker may result in a judgment that the sentence is somewhat unnatural. No such inference about an odd epistemic state is required for (1b), with the connective *and*. Here, the speaker simply knows (at least some) of the people who engaged in the activity. According to the epistemic state hypothesis, the harder it is to imagine a context in which the speaker would have some particular knowledge about an event but not other knowledge about that event, the less acceptable or natural an utterance implying that state will be --at least in an out of the blue context. However, if the context supports uneven knowledge of the event on the part of the speaker then the oddity of the utterance should be reduced or eliminated.

We begin by testing some implications of the QUD account of the phenomenon, with findings that seem to disconfirm it as the source of the phenomenon. We continue with tests of the epistemic state hypothesis, which provide positive support, and go on to explore the time course with which the need for accommodation is recognized. We conclude with an exploration of some of the possible range of applicability of the epistemic state hypothesis.

Experiment 1

The first experiment was a written questionnaire study, querying the coherence of two-person dialogs like (1).

Method

Subjects—Forty-four University of Massachusetts undergraduates participated in individual 15-minute sessions for course credit.

Materials—Twenty items like those in (1) were constructed, with two versions of each. (See Appendix 1 for all items.). The versions differed only in that in one form the sentence subject contained *or* while the other contained *and*. There were also ten filler items with subjects like (1a) but continuations that addressed (in 7 cases) which of the two possible entities was the agent (the remaining three had other coherent continuations), together with 4 practice items and 36 items from unrelated experiments (asking questions about the sequence of events or differences in interpretation of *how* and *when*, e.g., James put on his cap and tripped on a rock, followed by a question about whether this was one or two separate events, and How/When did Martha make up her mind? Martha made up her mind while going to the library.).

Procedures—Each subject was seated in front of a computer monitor that presented written instructions. They were told that some items would ask them to rate the coherence of 'Speaker B's' response to 'Speaker A' by typing a number from 1 to 7 on a keyboard, where 1 meant that Speaker B's sentence was "slightly off" and 7 meant it was "completely coherent." They also received instructions that other sentences would be followed by questions about the meaning of the sentences. The program Linger (Rohde, 2003) was used to present the practice items and then, in individually randomized order, the 20 experimental items intermixed with the 46 items from other experiments. A computer recorded the subjects' responses and reaction times (not analyzed).

Results

Table 1 presents the mean ratings (for Experiments 1-4), together with their standard errors. A linear mixed model (with random slopes and intercepts), implemented in the package LME4 (version 1.1-7) of R 3.2.1 (R Development Core Team, 2012) was used to analyze the data. It indicated that the effect of type of connective was fully significant ($b = 1.3$, $SE = 0.22$, $t = 6.00$) (throughout the paper, we take a $t > 1.96$ to be significant, since the t distribution approximates the z distribution for the large but unidentifiable df in a linear mixed model).

Discussion

Experiment 1 indicated that discourses that contained the disjunction *or* were judged to be less natural and coherent than matched discourses with *and*. It is tempting to ascribe this difficulty to how Speaker B's response failed to address the natural QUD suggested by the disjunction sentence, namely 'which of the two individuals mentioned engaged in the activity,' especially since subjects in the experiment were explicitly instructed to judge the

coherence of Speaker B's reply. However, subjects in experiments do not always follow instructions perfectly, so we cannot conclude that problems with the inferred QUD are behind the observed results. Experiment 2 was designed to test one implication of the proposal that failure of Speaker B to address the accommodated QUD was the root of the difficulty with the disjunction sentences.

Experiment 2

The second experiment addressed the question of whether the unnaturalness of dialogs with *or* would be reduced by providing an explicit QUD which Speaker B's response would address. Half the subjects in Experiment 2 read and judged the same sentences that were used in Experiment 1, while half read and judged them following a lead-in question like *Did anyone object to the new policy?* which was attributed to a third speaker. If the explicit QUD blocks the presumed QUD that a reader infers upon reading a disjunctive sentence like those in Experiment 1, then – assuming that Speaker B's response addresses this QUD – the difficulty of the disjunction sentences should be reduced or eliminated.

Method

Subjects—Forty-eight subjects were recruited over the internet. Half received lead-in questions (see Procedures) and half did not. All had ip addresses in the US, and reported that they were over 18 years old and had English as their native language. Each was paid \$1.00 for participation.

Materials—The 20 sentences used in Experiment 1 were used in Experiment 2, in both disjunctive and conjunctive versions. Two experimental conditions were added by prefacing each two-person discourse with a question attributed to a third speaker, as in (2):

(2) Speaker A: Did anyone object to the new policy?

Speaker B: Jeff or (and) Brayden objected to the new policy.

Speaker C: Dan did too.

Fifteen filler sentences were constructed with conjunctive or disjunctive sentences attributed to one speaker, and with plausibly responsive replies from a second speaker (e.g., identification of which of the two disjuncts mentioned by the first speaker actually engaged in the activity mentioned). The fillers were prepared with and without lead-in questions; the former were presented to subjects who received lead-in questions for the experimental items.

Procedures—Experiment 2 was conducted as a web-based experiment, using Amazon's Mechanical Turk (<https://requester.mturk.com/start>) to recruit subjects. Twenty-four subjects were assigned to materials like those in Experiment 1, with no lead-in question; 24 were assigned to the same materials with a lead-in question added. The experiment was presented to them using Ibex Farm (<http://spellout.net/ibexfarm>). They received instructions to rate the dialogs they would see in terms of how “normal and sensible” they were, using a 7-point scale where 7 was defined as fully sensible, something reasonable and cooperative speakers might say). After three practice items, they received the 20 experimental items and

15 fillers in individually-randomized order, indicating their rating after each sentence by typing a number or clicking a point on a 7-point scale. Two counterbalancing conditions were used, so that each sentence was tested equally often with *or* and with *and*.

Results

The mean ratings appear in Table 1. The results are very clear. The data were analyzed using a linear mixed model in R and LME4, as in Experiment 1. Random intercepts of subjects and items were used, and random slopes of the connective variable by subjects and items, but (because of the between-subject design) only random slopes by items of the lead-in question variable were used. The lead-in question and connective effects were treated as numeric variables, coded as ± 0.5 , so that each main effect of the analysis could be interpreted pooled over the other variable, as in ANOVA. The ratings were significantly lower for sentences with *or* than sentences with *and* ($b = 0.81$, $SE = 0.14$, $t = 5.67$). ($F(1, 46) = 32.23$, $p < .001$; $F(1, 19) = 172.17$, $p < .001$) and for dialogs with lead-in questions ($b = 0.49$, $SE = 0.24$, $t = 2.02$). The interaction was not significant ($b = -0.16$, $SE = 0.29$, $t = -0.57$).

Discussion

The prediction that the difficulty of disjunction sentences would be reduced or eliminated when subjects were provided with an explicit QUD that the final speaker's response addressed was not supported. The presence of the explicit question did improve ratings overall, but did not reduce the *or* penalty. While this finding does not conclusively demonstrate that there is no cost to the failure to address the QUD that is implicated by the disjunctive sentence, it certainly provides no support for the QUD hypothesis advanced earlier.

Experiment 3

The third experiment provided a second, very direct, test of the hypothesis that the problems with disjunctive sentences stemmed from the failure of the second speaker's response to address the QUD. It simply compared ratings when there was and was not such a second speaker's response. To the extent that the problems with the disjunctive sentences remain when there is no response, they cannot be attributed to a failure of the response to address the QUD implicated by the sentence with the disjunction.

Method

Subjects—Thirty-two subjects with the same characteristics as in Experiment 2 were recruited over the internet and paid \$1.00 for participation.

Materials—The same basic materials used in Experiments 1 and 2 (without the lead-in sentences of Experiment 2) were used in Experiment 3, together with the fillers of Experiment 2. Besides eliminating Speaker B's response on half the trials, one change was made in Experiment 3: the disjunctive sentences were introduced by *either* (e.g., *Either Jeff or Brayden objected to the new policy*) and the conjunctive sentences by *both*. We thought that blocking the inclusive reading of the disjunctive phrase might increase its naturalness.

Two versions of each discourse were made, one with “Speaker A” and “Speaker B” where “Speaker B” responded to “Speaker A” as in the first two experiments, and one version with no “Speaker B” response.

Procedures—The procedures used in Experiment 2 were used in Experiment 3, apart from the fact that the manipulation of both *and* vs. *or* and presence vs. absence of a “Speaker B” response were manipulated within-subjects. The subjects were recruited using Mechanical Turk as in Experiment 2, and tested using the same procedures used in that experiment.

Results

The mean ratings appear in Table 1. The data were analyzed using a linear mixed effect model with random slopes and intercepts, as in Experiment 1. Numeric coding of the fixed effects was used. Once again, disjunctive sentences received lower ratings than conjunctive sentences ($b = -1.41$, $SE = 0.17$, $t = 8.19$). Further, two-sentence dialogs received lower ratings than single sentences ($b = -0.68$, $SE = 0.11$, $t = 5/97$). The interaction was nonsignificant ($b = -0.37$, $SE = 0.26$, $t = -1.43$). The cost of *or* was numerically somewhat larger than that of *and*, leaving open the possibility that some of the low naturalness ratings of disjunctive discourses is due to QUD problems.

Discussion

Disjunctive sentences received lower naturalness ratings than conjunctive sentences, even in the absence of a reply that fails to address the presumed QUD that readers accommodate for a disjunctive sentence. Finding a significant drop in acceptability without a continuation sentence shows that the *or* sentences have some complexity all by themselves, independent of the appropriateness of a response to the QUD that they may introduce. While we cannot reject the hypothesis that there is some effect of a failure to address the QUD that the disjunctive sentence presumably introduces, the experiment does allow us to conclude that the disjunction carries some cost, even in the absence of a reply.

This cost is unlikely to be due to the presence of two possible interpretations for *or*, since the presence of *either* strongly biases *or* to an exclusive interpretation. In Experiment 3 the penalty for disjunction should have been eliminated if the ambiguity of *or* were its source. Experiment 4 was designed to test the alternate hypothesis that was presented in the Introduction, the hypothesis that the difficulty of sentences with disjunctive subjects stems from the odd epistemic state of a speaker who knows that at least one (or both) of two people engaged in an activity, but does not know which one. Experiment 4 also provides a test of another hypothesis, suggested by an anonymous reviewer. It is possible that the dispreference for an isolated disjunction sentence stems from its failure to provide a complete answer to the QUD it is taken to address, namely, “Who left?”, while an (exhaustified) conjunction sentence does provide a complete answer. We will refer to this hypothesis as the ‘answer incompleteness’ hypothesis.

Experiment 4

The fourth experiment had two goals: first, to see if the persistence of the penalty for disjunctive sentences without a reply could be replicated in the absence of *either/both*, and

second, to provide a test of the epistemic state hypothesis presented earlier. This hypothesis suggested that sentences containing *or* might be complex simply because the speaker must be assumed to be in a peculiar knowledge state in order to produce a sentence with a disjunctive subject. The speaker must know that one of two people performed the activity described in the sentence, but know which one. While this is a possible state, by hypothesis the information needed to accommodate it goes beyond that which the speaker should assume is easily available, and the resulting difficulty of accommodation could lower the perceived naturalness of a disjunctive sentence.

The epistemic state hypothesis can be formulated as follows: Sentences which would be naturally uttered only in an unusual epistemic state are rated lower than those implying a more common epistemic state unless the discourse contains information supporting the accommodation of the relevant circumstances.

We focus on one prediction of this hypothesis (though see the General Discussion for further predictions). As asserted above, it is not particularly common for a speaker to know about an event that it happened, not know the identity of the agent (or experiencer), but know that the agent was one of two individuals. However, for sentences containing the future, it is much easier to imagine that a speaker might know that someone will perform the activity, without knowing which of two individuals it will be. In many circumstances, the same knowledge that would constitute evidence that some state of affairs will hold in the future also explains why there is only partial knowledge of the agent, e.g., the knowledge that one or the other or possibly both of the individuals served as the agent. In short, knowledge of future states is commonly partial whereas knowledge of past events, we hypothesize, typically includes knowledge of the identity of the agent (or the experiencer for other types of eventualities).

To test the epistemic state hypothesis we contrasted past and future versions of the sentences used in Experiment 3 (plus 4 additional sentences; see the Appendix). The epistemic state hypothesis leads to the expectation that there will be a larger penalty for disjunction sentences in the past than in the future. As in the other experiments, we expect lower acceptability for disjunction than conjunction, and as in Experiment 3, we expect the penalty to hold regardless of the presence of a continuation. The experiment was conducted twice, and the data will be reported as Experiment 4A and 4B.

This experiment also provides a test of the answer incompleteness hypothesis described in the Discussion of Experiment 3, namely, that isolated disjunctive sentence in contrast to isolated conjunctive sentences provide only a partial answer to the QUD they may be presumed to address, namely, "Who left?". Partial answers may be judged less acceptable than complete answers. However, the contrast in completeness *per se* between conjunctive and disjunctive sentences holds equally for future and past tense sentences, and thus no effect of tense is predicted by the hypothesis that the difficulty with the disjunctive phrase is just that it is a partial answer to the presumed QUD. Hybrid accounts are of course possible, e.g., a partial answer may be more acceptable when partial knowledge is easier to countenance, as in the future, but the success of such accounts does seem to rest on a principle very similar to the core principle of our epistemic account.

Method

Subjects—For Experiment 4A, 40 subjects were recruited through Mechanical Turk. For Experiment 4B, a different 48 subjects were recruited. Each was paid \$1.00.

Materials—Eight versions of each of the 20 sentences used in Experiment 3 plus an additional four sentences were constructed by adding future tense versions to the past tense sentences used in that experiment (e.g., *Jeff or Brayden will object to the new policy.*). Again, there were versions with *or* and *and*, and versions with and without a continuation. The 15 filler discourses used in Experiment 3 were used in Experiment 4, with six changed from past to future tense.

Procedures—Both Experiment 4A and 4B were conducted using the same procedures as used in Experiments 2 and 3. Eight different counterbalancing conditions were used, five (Experiment 4A) or six (Experiment 4B) subjects per condition, to ensure that each item was tested equally often in each of the eight forms of the experiment. Because of a coding error, four items in Experiment 4A were tested in only some of the 8 experimental conditions. This problem was corrected in Experiment 4B.

Results

The mean ratings for both experiments appear in Table 1. They were analyzed as in Experiment 3, using a linear mixed model with fixed effects of $\pm$ continuation, *and/or*, and past/future. Numeric coding of the fixed effect factors ($\pm$ 0.5) was used, permitting ANOVA-style interpretation of the main effects. In Experiment 4A, the maximal model, with full random slopes and intercepts by subjects and items, did not converge. Convergence was obtained by eliminating random items slope of the factor of tense (which exploratory analysis suggested contributed the smallest variance to the random effects). In Experiment 4B, the maximal model converged, and its results will be reported.

The results are very simple. The familiar penalty of disjunctive sentences was observed (Experiment 4A: $b = 1.32$, $SE = 0.20$, $t = 6.62$; Experiment 4B: $b = 1.28$, $SE = 0.19$, $t = 6.87$) and the interaction of tense and connective was significant (Experiment 4A: $b = -.63$, $SE = 0.13$, $t = -4.96$; Experiment 4B: $b = 0.37$, $SE = 0.11$, $t = -3.34$). The disjunction penalty was significantly smaller for future tense sentences than for past tense sentences (although it was significant for each: Experiment 4A: future, $b = 1.00$, $SE = 0.22$, $t = 4.69$; past, $b = 1.63$, $SE = 0.22$, $t = 7.36$; Experiment 4B: future, $b = 1.09$, $SE = 0.20$, $t = 5.63$; past, $b = 1.47$, $SE = 0.21$, $t = 7.02$). In Experiment 4A, the small overall advantage for future over past (5.59 vs 5.34) was significant ($b = 0.25$, $SE = 0.07$, $t = 3.70$); this effect became nonsignificant in Experiment 4B ($b = 0.00$, $SE = 0.08$, $t = 0.04$). No other effects reached significance.

Discussion

Experiment 4 provides evidence for the epistemic state hypothesis: the cost of a disjunctive subject was reduced for future tense sentences, arguably because using the future tense made it easier for the reader to accommodate the apparent mental state of the source of the sentence. It is, quite simply, more likely for someone to believe that one of two people *will*

perform an action than to know that it is one of these two who performed the action already but to not know which one. The effect of tense also provides an argument against the answer incompleteness hypothesis, at least as a complete account of the phenomenon: the future tense answer is as incomplete as the past tense answer. There is, however, still some penalty to the disjunctive subject in future tense sentences. It may be more natural to assume that a speaker knows that two people will perform some action in the future than to know that one or both of the two will perform it. In any case, the presence of a penalty in the future suggests accommodation of an uncomfortable epistemic state may not be the only source of difficulty with the disjunctive sentence.

Experiment 4, however, does provide more evidence that failure of a reply to address the QUD that is presumably accommodated for a disjunction (“Which of the two did the action?”) is not a major factor in the apparent difficulty of disjunctive subjects. As in Experiment 3, this difficulty was not reduced by removing the reply, showing that the reply per se was not the source of the problem.

The final experiment moves away from off-line naturalness judgments to study the time course of the apparent difficulty of disjunctive sentences. It uses eyetracking while reading to see whether there is an on-line cost of disjunction (as opposed, e.g., to some unidentified, perhaps prescriptive, preference to avoid disjunctions). It also provides further tests of the epistemic state hypothesis and the QUD hypothesis.

Experiment 5

The final experiment asked whether the off-line penalty for the disjunctive discourses seen in Experiments 1 - 4 would appear as an on-line effect during normal reading. The epistemic state hypothesis predicts that the disjunctive penalty could appear on the disjunctive phrase itself in the first sentence (as well as possibly later in the discourse, if the reader persists in having difficulty accommodating the disjunction). In contrast, the QUD hypothesis predicts that the disruption will not appear until the second sentence -- the response that by hypothesis fails to address the QUD implicitly introduced by the disjunction. There is reason to expect that QUD-based disruption would appear early in the second sentence, at the point where the relevant information is introduced. Clifton and Frazier (2012) found that readers' eye movements were disrupted quickly when they read material that failed to address the QUD suggested by earlier material (the verb bias of a question, in their Experiment 1; the presence of a modal that introduced uncertainty, in their experiments 2 and 3). The disruption appeared in the form of more frequent regressions from a region if it failed to address the assumed QUD than if it did address it. In the present disjunction dialogs, this would suggest quick disruption of reading of Speaker B's response to a disjunction sentence, if the difficulty is that the response failed to address the QUD accommodated to deal with the disjunction. However, if (as described earlier), the difficulty of disjunction sentences reflects the difficulty in accommodating the mental state of the person who used *or*, any difficulty may well appear on the first sentence. Further, under this epistemic hypothesis, if the reader delays accommodation to see if any relevant information appears late in the discourse, any disruption in the second sentence may appear only late in the sentence.

Experiment 5 also contained a novel test of the epistemic state hypothesis. Half of the dialogs were like those used so far, but half were introduced by an indirect question. In these versions, Speaker A's sentence began *I wonder if it was....*, suggesting that the speaker was in a state of uncertainty about who performed the action. This uncertainty extended to both the disjunction and conjunction sentences, but could reasonably be expected to affect primarily the disjunction sentences. The speaker is explicitly saying that s/he is in the epistemic state of believing that one of two people did some action, but not knowing which one. This is the state which we suggested readers of the disjunctive dialogs in earlier experiments would have to accommodate, and its explicit statement could very well relieve the reader of the difficulty of accommodation.

The presence of the *I wonder if it was...* lead-in also makes the QUD overt. Speaker B's response, to the effect that an additional person engaged in the behavior, is clearly unresponsive to this QUD. According to the QUD hypothesis, this should thus make the reply to the disjunctive sentence worse, not better, than its no-lead-in counterpart. In sum, the epistemic state hypothesis leads to the expectation that the effect of the lead-in should reduce the disjunctive penalty and it could appear early (in sentence one) or perhaps also at the very end of the discourse. The QUD hypothesis leads to the expectation that the effect of the lead in should increase the disjunctive penalty, and should appear early in Speaker B's reply.

Method

Subjects—A total of 70 University of Massachusetts undergraduates contributed data to be presented. An additional 23 were tested but their data were rejected for having eyetracking problems on more than 25% of the experimental sentences in a preliminary analysis described below. The subjects participated for course credit.

Materials—Versions of the 20 two-person dialogs used in Experiment 1 were used in Experiment 5, with the Speaker B response extended by a few words so that one possible critical region (the 'did too' response) would not appear at the end of a sentence (where eye movements may be disrupted in various ways). Further, two new versions of each sentence were made up, with a lead-in phrase. An example appears in (3), where the lead-in phrase appears in parentheses, and the regions of the sentences to be analyzed are indicated by numerical subscripts. All materials appear in the Appendix.

(3) Speaker A: /₁(I wonder if it was)/₂ Jeff or/and Brayden/₃ (who) objected to the new policy./₄

Speaker B: /₅Dan did too, but it didn't do/₆ any good./₇

As in the previous experiments, two versions of the items contained a conjunction and two conditions a disjunction in the subject phrase. Orthogonally, two versions contained a clefted indirect question lead-in, which could sharpen the reader's accommodation of a QUD like "Which one?" or could indicate the epistemic state of Speaker A. These items were combined with the 10 disjunctive filler dialogs used in Experiment 1, together with 8 practice items and 90 other sentences or/and dialogs from unrelated experiments (examining

the processing of relative clauses, focus effects, etc.; examples appear in the Appendix). Comprehension questions were made up for a total of 48 of these 128 items (the questions for the current experiment appear in the Appendix).

Procedures: Subjects' eye movements were recorded while they read the experimental sentences for comprehension, using an Eyelink 1000 eye-tracker (SR Research, Toronto, Ontario, Canada), connected to a PC computer, sampling at 1000 Hz. The subjects viewed the two-line discourses binocularly on a CRT monitor from a distance of 55 cm. They were always instructed to read sentences (presented in 11-pt Monaco font) silently and at their own pace, and they could take as many breaks as they wanted. The experiment was controlled by the Eyetrack software (<http://blogs.umass.edu/eyelab/software>).

At the beginning of the experiment, all subjects were presented with a 13-point grid in order to calibrate their eye movements with the locations on the screen. If the initial calibration was especially difficult due to factors such as glare from glasses, the experimenter presented the subjects with a 9-point calibration grid instead. Recalibration between trials took place as needed. Each trial was triggered when the subject fixated a box on the left edge of the computer screen.

Before seeing the target stimuli and the fillers, participants were presented with seven 2-line practice trials. The experimental items were then presented along with the fillers in a pseudo-randomized order. "Speaker A's" assertion and "Speaker B's" response appeared simultaneously on the monitor, one above the other, each labeled as "SPEAKER A:" and "SPEAKER B:". Each subject only saw one possible condition for each item; four counterbalancing conditions were used to ensure that each sentence was seen equally often in each experimental condition. To encourage careful reading, comprehension questions were asked after four of the practice items and 10 of the experimental items.

Data analysis: A preliminary examination of the data was made using the program EyeDoctor (<http://blogs.umass.edu/eyelab/software/>) to correct vertically-misaligned fixations and to identify trials on which there was a track loss (caused often by an eyeblink) in one critical region of the sentence, the "...did too" region, region 6 in (12). 23 subjects' data were discarded for losing more than 5 trials because of track losses. Of the remaining subjects, 8 % of the trials were lost for this reason.

The data were then analyzed (using the program EyeDry, available at <http://blogs.umass.edu/eyelab/software/>). Several conventional eyetracking measures (Rayner et al, 1989) were used, including first pass time (the summed fixation duration from first entering a region until first leaving it), total time (the sum of all fixation durations in the region), go-past time (the sum of all fixation durations from first entering a region until leaving it to the right, and proportion of first pass regressions out of a region. The sentences were segmented as indicated by subscripts in (2). Regions 3, 4, 6 and 7 are the regions of interest. Region 3 contained the conjunctive or disjunctive phrase, and Region 4 was the remainder of the first sentence. Region 6 contained the elliptical phrase plus the following three words; Region 7 contained the few words at the end of the sentence.¹ The results (in

mean durations, ms, or/and proportion of regressions out), together with standard errors (in parentheses) appear in Table 2.

Significance of any effects was evaluated using linear mixed models in R Version 3.2.1 (logistic models, in the case of proportion of regressions out) with fixed effects of type of connective (*and* vs *or*) and presence or absence of a clefting lead-in to the sentence. As in the previous experiments, the fixed effects were coded as numerical variables, weighted as ± 0.5 , so that the effect of each factor was tested over the pooled levels of the other factor as in a conventional ANOVA. Interacting random slopes were used, except when the analysis failed to converge, in which case non-interacting random slopes were used (except in the case of analyses of regressions out, where the simplified models that converged are specified). As before, a *t* value of 1.96 or/and greater was treated as significant, assuming that *t* approximated *z* given the large but uncertain *df* in the models.

The core results, presented in Table 2, are quite straightforward. There was evidence of slow and disrupted initial reading of the initial sentence when it contained a disjunction, compared to a conjunction, but only in the absence of the "I wonder if..." (clefting) lead-in. This reflected the lower acceptability of simple disjunctive sentences, as demonstrated in the preceding experiments. There was also evidence of a delayed disruption in reading the reply to a disjunctive sentence, disruption that appeared largely in measures that are sensitive to time spent re-reading material as opposed to initially reading it and appeared for all discourses with disjunctions, regardless of the lead-in..

First pass time on the phrase with the connective (Region 3 in example (3)) was faster in the presence of the clefting lead-in than in its absence ($b = 117.64$, $SE = 17.15$, $t = -6.86$). This probably reflects the possibility of a quick regression from Region 3 back to the lead-in (see Rayner & Sereno, 1994, for discussion), regressions which were unlikely in the absence of a lead-in. More critically, there was a marginally-significant interaction with type of connective ($b = 49.45$, $SE = 26.08$, $t = 1.90$). When there was no lead-in, the disjunctive phrase was read more slowly than the conjunctive phrase ($b = 48.51$, $SE = 20.45$, $t = -2.37$). This difference disappeared with the lead-in phrase ($b = -3.11$, $SE = 20.45$, $t = -0.15$). The significant effect of the lead-in can be attributed largely to the slow reading of a disjunctive phrase when the lead-in was absent. The same interaction, now fully significant, was observed in go-past time ($b = 84.75$, $SE = 30.46$, $t = 2.78$). The simple effect of lead-in became non-significant in this measure (which includes time re-reading the lead-in) ($b = 30.08$, $SE = 19.36$, $t = 1.55$). The interaction was also significant in total time ($b = 121.45$, $SE = 44.01$, $t = 2.76$), as was the effect of presence of lead-in ($b = 112.06$, $SE = 29.17$, $t = 3.84$, largely reflecting the contribution of the slower first-pass times when the lead-in was absent). In each case, a simple effects test indicated that the disjunctions were read more slowly than conjunctions when the lead-in was absent (go-past $b = 55.10$, $SE = 21.67$, $t = 0.254$; total time $b = 92.28$, $SE = 37.21$, $t = 2.48$). The difference was nonsignificant when the lead-in was present (maximum $t = 1.23$). Analyses of regressions out (which converged only with random intercepts, not random slopes) indicated that they were more frequent

¹Region 6 was also analyzed as two separate regions, the region containing the elliptical phrase and the remaining words. The patterns of results were the same in both regions, but never reached statistical significance when the regions were considered separately.

when a lead-in was present ($b = -1.91$, $SE = 0.28$, $z = 6.88$, $p < .001$), which can largely be attributed to the fact that there was substantive material to regress into in this case.

Analyses of Region 4, the remainder of the first sentence, were less informative. No effects reached significance in the analyses of go past or total time. The interaction between lead-in phrase and type of connective approached significance in the analysis of first pass time ($b = -107.0$, $SE = 62.27$, $t = -1.72$), perhaps reflecting the significant interaction observed in frequency of regressions out ($b = 0.99$, $SE = 0.32$, $z = 3.11$, $p < .01$; random slopes of connective had to be eliminated to make the model converge). Readers appeared to remain fixated in Region 4 in the *or*, lead-in condition, while they tended to regress more frequently in the other conditions.

Next, consider the time to read the second sentence, containing the *did too* response. A significant impairment of reading responses to initial disjunctive (vs. conjunctive) sentences was found, but only in late measures. The clearest effects appeared in the total time measure, where replies to disjunction sentences were read more slowly than replies to conjunction sentences Region 6 ($b = 49.28$, $SE = 23.30$, $t = 2.12$). Regressions out of Region 7 were marginally more frequent following *or* than *and* ($b = 0.159$, $SE = .0086$, $t = 1.84$). All in all, there is evidence for disruption following *or* toward the end of the response sentence, and appearing primarily as re-reading of the early part of the sentence. There was little or no evidence for any slow-down in initial reading of the sentence. Reading of Region 6 was also slower overall when Speaker A's sentence was clefted (with a lead-in) than when it was a simple sentence. The significant effects appeared on Region 6 for go-past time ($b = -82.45$, $SE = 31.35$, $t = 2.63$) and total time ($b = -85.94$, $SE = 24.42$, $t = 3.52$). However, there were no interactions with type of connective.

Discussion

The results of the eye movement study provide some support for the epistemic state hypothesis, but very little support for the QUD hypothesis as it has been advanced. Reading time of the disjunctive phrase was slowed when it appeared on its own, but this slowing was reduced or eliminated in the presence of the lead-in phrase. As suggested earlier, this lead-in is an indirect question which could very plausibly convey Speaker A's uncertainty to the reader, justifying use of *or*. The lead-in might have been expected to serve as an explicit QUD, affecting the time to read the *did too* reply, but any effect it had there was the same for sentences with *and* and *or*. This suggests that the late disruption for *or* sentences may not reflect the same difficulty as that observed on the disjunctive phrase itself, which we propose resulted from problems in accommodating the epistemic state of a speaker who used the disjunction. Thus, it is still possible that there is some residual difficulty caused by a reply that fails to address the presumed QUD, but we cannot claim to have any clear evidence for or against this.

General Discussion

To summarize, the current experiments showed that discourses containing the connective *or* were generally judged less natural, and processed with greater difficulty, than matched discourses containing *and*. No evidence for the QUD hypothesis was obtained. There was no

clear evidence that the difficulty of *or* lay in the failure of a response to address the presumed QUD *which one did it?*, no evidence that providing an appropriate explicit QUD relieved the difficulty of *or*, and no evidence of the sort of quick disruption of reading a continuation that has previously been observed for material that fails to address the likely QUD. We do not conclude that claims that the QUD guides discourse comprehension have no merit. On the contrary, we do believe that such claims have a great deal of merit. We simply assert that there is no reason to believe that the reply's failure to address a likely QUD is at the root of the phenomenon we have explored here.

In contrast, some evidence was presented in support of the hypothesis that the difficulty of *or* reflected the difficulty of accommodating an epistemic state that would permit "Speaker A" to produce a disjunctive sentence. The cost of the disjunction was reduced for future-tense sentences in Experiment 4. Further, in Experiment 5, the difficulty of reading a disjunctive phrase appeared to disappear in the context of a lead-in that suggested that the speaker might be in an appropriate epistemic state.

In principle, the cost of disjunction might be due to the complexity of the corresponding mental model representation (Goodwin & Johnson-Laird, 2011). When there is a disjunction, more than one possibility must be represented and this may come at a cost in terms of representational complexity. However, the fact that the cost of the disjunction was mitigated when the *I wonder if it was* lead in was present does not fit well with this approach since the representational complexity of the disjunction should be present regardless of the preceding context. Further, a representational complexity account would not explain the greater cost for disjunction in the past than in the future.

Instead we have attributed the penalty for disjunction to the oddity, in an out of the blue or unsupportive context, of the listener assuming that the speaker knows that an event took place and that it had one of two particular individuals as its agent, but not know which one.² This account leads one to expect that the context *I wonder if it was X or Y* would mitigate the cost of disjunction because the speaker's epistemic state is not accommodated but instead is at issue, and that the cost of disjunction would be greater in past tense contexts, where this type of partial knowledge of an event seems harder to imagine, than in future contexts, where partial knowledge is the norm. Typically listeners and readers readily accommodate unstated information, as we stressed in the introduction, provided that the information is salient in the discourse context or the information involves shared beliefs of the interlocutors that can be assumed to be expected and uncontroversial.

The epistemic state hypothesis suggests additional predictions. It may be easy to imagine a context in which a speaker is fully knowledgeable about a situation she is describing but less easy to imagine certain uneven or incomplete epistemic states, e.g., knowing that some event took place before time X at Location A or it took place after time Y at Location B. But once given a context giving rise to the uneven knowledge, the difficulty or acceptability penalty

²One might imagine that the penalty for disjunction is limited to subjects containing a disjunction, as in the present experiments. However, that is not the case. In an unpublished study largely comparable to Experiment 1 but with the disjunction in either the subject or the predicate (e.g., Jennifer criticized Jeff or Brayden at the meeting), very similar disjunction penalties were observed (a penalty of 0.75 points on a 7 point scale for subject disjuncts, a penalty of 0.91 points for predicate disjuncts).

should be ameliorated. Indeed, as noted by a reviewer, in the context of a detective game where the players collectively figure out who left, *John or Bill left. Sam did too* improves. If unsupported partial knowledge is the source of difficulty in disjunction, as we have claimed, then various signals of that partial knowledge state should also improve the status of the examples. These might include conditionals which implicate speaker ignorance of both the antecedent and consequent, as pointed out by a reviewer (4), adverbs such as *apparently in* (5), and explicit disavowals of knowledge (5).

(4) a. If John or Bill left, Sam did too.

b. If John and Bill left, Sam did too.

(5) Apparently John or Bill left. Sam did too.

(6) John or Bill left, though I don't know which. Sam did too.

It is commonplace to talk about real world plausibility in studies of language processing. Language processing times are expected to be slower when unlikely situations are described, (e.g., *We ate dinner under the table*), than when likely situations are described (*We ate dinner at the table*; Rayner, Warren, Juhasz, & Livens, 2004). Epistemic knowledge effects may be seen as another type of plausibility. Likely epistemic states do not need contextual justification whereas unlikely ones do. Viewed this way, the present results might be related to those in Van Berkum et al. (2008). In ERP studies when a speaker and a message mismatched, a small but reliable N400 was observed -- e.g. when a female voice said *I always rent movies with lots of violence* or a voice in an upper class accent said *I have a big tattoo on my back*.

In the present studies, the penalty for disjunction was observed both in acceptability judgments and in reading times. The acceptability judgment effects observed in Experiments 1-4 are informative with respect to how acceptability judgments are formed. In principle, judgments might be influenced only by the grammatical well-formedness of the utterance and the complexity of assigning a grammatical structure and meaning to it. But that is not what we have observed. Instead, if a sentence implies an unusual epistemic state this can lower the rated acceptability of a sentence. Put differently, if a listener can easily imagine a circumstance in which a sentence might be uttered, then the sentence is judged somewhat more acceptable than if not. In prior work involving English sentences lacking phonologically overt subjects (Mack, Clifton, Frazier & Taylor, 2012), we have seen evidence that dovetails with this conclusion. In particular, present tense utterances like (*It seems cold to me*) were more likely than equivalent past tense sentences to be heard in noise as lacking a subject, and in acceptability judgments subjectless present tense sentences were rated as more acceptable than past tense counterparts, even with nonsense verbs (Frazier, 2015, in press). We attributed these findings to the use of null subject sentences (*seems cold to me*) to comment on perceptually available states of affairs, which itself correlates with present tense grammatical features. Because of the large role that acceptability judgments play in linguistic theorizing, it is important to recognize the various kinds of factors that influence acceptability judgments. The likely epistemic state of the speaker, and the

circumstances in which a sentence might be used, seem to be among the factors reflected in such judgments.

Turning to the eye movement data, the interaction of connective and context is highly informative. One might imagine that various types of linguistic and general world information might be exploited by the reader to arrive at a propositional interpretation of a clause and then the reader might draw further inferences about the speaker/author's epistemic state and her intent in uttering the clause. But the finding that early reading times on *or* itself, including first pass reading times, show a significant interaction of connective and context, with shorter times when a *I wonder if* lead in is present, suggests that the input is immediately related to the speaker's knowledge state. Many researchers, especially those working in the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), have emphasized the on-going relation between language processing and the external context of an utterance, especially when the listener must operate on that context (e.g., *touch the ...*). The present findings suggest that language processing is being related incrementally to the speaker/author's epistemic state even in the absence of an explicitly given context.

Although we have used disjunction sentences to bring forward the effect of epistemic states, we do not think epistemic state penalties are tied always or only to disjunction. Consider (7).

- (7) a. John and Bill didn't go to the grocery store.
b. John didn't go to the grocery store.

The conjunction sentence in (7a) seems odd unless one accommodates the assumption that John and Bill were expected to go to the grocery store. Intuitively, it is odder than *John didn't go to the grocery store* perhaps because additional unsupported accommodation is needed for the conjunction sentence (John and Bill separately, collectively) relative to the sentence with a simple subject (7b).

Although we suspect that epistemic state effects are pervasive and that we have only scratched the surface of their role in language comprehension, we have provided evidence that they do affect acceptability ratings and online comprehension. Epistemic effects are not only exquisitely sensitive to the very particular content of sentences, but they show very general effects as well. Consider (8).

- (8) a. Everyone or no one came.
b. Everyone or no one will come.

In the past tense (8a), the sentence seems downright odd. How could the speaker have evidence that one of the two states described obtained? But now consider its future counterpart (8b). One can rather easily imagine a nervous hostess before a party concerned about the amount of food, worrying both about whether it is sufficient and about whether it is too much. The sizeable difference in intuitions about (8a) and (8b) attests to the importance of the speaker's (inferred) epistemic state in language comprehension. Among

other things, we have presented evidence here that the past and the future have different affordances for assessing the speaker/author's epistemic state.

A sentence might be puzzling in terms of the speaker's epistemic state at the beginning of the utterance but made clear by information later in the utterance. For example, (9a) seems slightly odd. It's not quite clear what the speaker's epistemic state is.

(9) a. John or Bill didn't go camping.

b. John or Bill didn't go camping. I saw one of their pup tents in the garage.

But (9b) seems much more natural. The continuation sentence provides the information needed to figure out what the speaker's epistemic state is. In (9b) less unsupported or free form accommodation is required. We suspect that the reason for the late reading time penalty for disjunction discourses in the eye tracking study was that much of the penalty due to an uncomfortable or unusual epistemic state arises only once the discourse ends without providing support for whatever unusual accommodation is needed.

In sum, listeners and readers seem to infer the speaker's epistemic state as they process a sentence incrementally with some processing cost when it appears that the epistemic state involves unmotivated uneven partial knowledge of the states of affairs described. But it appears that readers also hold out the possibility that later material will rescue a sentence, implicitly explaining how the speaker might come to have partial knowledge of the state of affairs described. Although the effect of an uncomfortable epistemic state is somewhat subtle intuitively, it influences acceptability judgments in ways that are reliable within and across experiments, and it influences online processing in ways that suggest both an immediate and a later cost. That such effects should exist is perhaps not surprising given the important role of accommodation in ordinary language comprehension. Accommodation of presuppositions and of other unstated information plays a critical role in comprehension and this in turn makes consistency in the epistemic state of the speaker or author paramount in language comprehension.

Acknowledgment

This project was supported in part by Grant Number HD18708 from NICHD to the University of Massachusetts. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of NICHD or NIH.

Appendix

Materials used in Experiments 1-4 (with modifications described in text). Note, sentences with *or/and* and with *and* combined as *and/or*

1. SPEAKER A: Jeff and/or Brayden objected to the new policy. SPEAKER B: Dan did too.
2. SPEAKER A: Arianna and/or Chelsea went to Paris this summer. SPEAKER B: Sarah did too.

3. SPEAKER A: Lindsay and/or Caitlin mentioned the High school's peeling paint.
SPEAKER B: Kelly did too.
4. SPEAKER A: Michael and/or Ian dropped off some food for the party. SPEAKER B: Kevin did too.
5. SPEAKER A: Angie and/or Katie photographed the sunset at Look Park. SPEAKER B: Phil did too.
6. SPEAKER A: Evan and/or Andrew called the boss about the obnoxious coworker.
SPEAKER B: Phil did too.
7. SPEAKER A: Marie and/or Jasmina walked to Amherst center. SPEAKER B: Meghan did too.
8. SPEAKER A: Adam and/or Russell lifted weights at the gym. SPEAKER B: Jack did too.
9. SPEAKER A: Kyle and/or John drank three cups of coffee. SPEAKER B: Seth did too.
10. SPEAKER A: Jess and/or Laura watched Humphrey Bogart films all weekend.
SPEAKER B: Gina did too.
11. SPEAKER A: Sam and/or Cory rejected insurance claims. SPEAKER B: Pat did too.
12. SPEAKER A: Amy and/or Andria worked on a science experiment. SPEAKER B: Kira did too.
13. SPEAKER A: Sophie and/or Emily grew beautiful tomatoes. SPEAKER B: Kira did too.
14. SPEAKER A: Josh and/or Tim ate pasta every night. SPEAKER B: Ryan did too.
15. SPEAKER A: Andrew and/or Liam always wore a kilt. SPEAKER B: Duncan did too.
16. SPEAKER A: Jeremy and/or Spencer went to law school. SPEAKER B: James did too.
17. SPEAKER A: Nicholas and/or Dan played Angry Birds. SPEAKER B: Aiden did too.
18. SPEAKER A: Claire and/or Liz spoke up against the new regulations. SPEAKER B: Minta did too.
19. SPEAKER A: Kara and/or Emma shopped at Big Y. SPEAKER B: Katie did too.
20. SPEAKER A: Gary and/or Tom fed the dogs. SPEAKER B: Ryan did too.
- ADDITIONAL ITEMS USED IN EXPERIMENT 4
21. SPEAKER A: Mary and/or Sue brought some snacks. SPEAKER B: Betsy did too.
22. SPEAKER A: Phil and/or Kyle ran in the Marathon. SPEAKER B: Pete did too.

23. SPEAKER A: James and/or Janice brought some wine to the party. SPEAKER B: Phyllis did too.

24. SPEAKER A: Frieda and/or Samantha tried on some new dresses. SPEAKER B: Sara did too.

Materials used in Experiment 5, collapsed over *or/and*, with optional lead-in phrase (*who* was deleted without lead-in); comprehension questions and example filler items are also presented

1. SPEAKER A: ((I wonder if it was)) Jeff or/and Brayden who objected to the new policy.

SPEAKER B: Dan did too, but it didn't do any good.

Did Dan get the policy changed? YES NO

2. SPEAKER A: ((I wonder if it was)) Arianna or/and Chelsea who went to Paris this summer.

SPEAKER B: Sarah did too, and enjoyed it immensely. What did Sara enjoy? Rome Paris

3. SPEAKER A: (I wonder if it was) Lindsay or/and Caitlin who mentioned the High school's peeling paint. SPEAKER B: Kelly did too, and said it was yucky.

What was the high school like? Old and run-down New and shiny

4. SPEAKER A: (I wonder if it was) Michael or/and Ian who dropped off some food for the party. SPEAKER B: Kevin did too, and people were grateful.

What was the food for? A food pantry A party

5. SPEAKER A: (I wonder if it was) Angie or/and Katie who photographed the sunset at Look Park. SPEAKER B: Christine did too, and the result was beautiful.

Is Christine a good photographer? YES NO

6. SPEAKER A: (I wonder if it was) Evan or/and Andrew who called the boss about the obnoxious coworker. SPEAKER B: Phil did too, but he couldn't get anything done either. Did Phil get the coworker fired? YES NO

7. SPEAKER A: (I wonder if it was) Marie or/and Jasmina who walked to Amherst center. SPEAKER B: Meghan did too, and thought it was beautiful.

Where did Marie and Jasmine walk? Amherst Northampton

8. SPEAKER A: (I wonder if it was) Adam or/and Russell who lifted weights at the gym. SPEAKER B: Jack did too, and worked up a real sweat.

What did Russell do at the gym? Play basketball Lift weights

9. SPEAKER A: (I wonder if it was) Kyle or/and John who drank three cups of coffee.
SPEAKER B: Seth did too, and felt pretty wired.

What did Kyle drink?.Coffee Power drink

10. SPEAKER A: (I wonder if it was) Jess or/and Laura who watched Humphrey Bogart films all weekend. SPEAKER B: Gina did too, and loved them all.

What did Laura watch? TV Movies

11. SPEAKER A: (I wonder if it was) Sam or/and Cory who rejected insurance claims.
SPEAKER B: Pat did too, disappointing her clients.

12. SPEAKER A: (I wonder if it was) Amy or/and Andria who worked on a science experiment. SPEAKER B: Molly did too, hoping to go to the science fair.

13. SPEAKER A: (I wonder if it was) Sophie or/and Emily who grew beautiful tomatoes.
SPEAKER B: Kira did too, and gave them to her neighbors.

14. SPEAKER A: (I wonder if it was) Josh or/and Tim who ate pasta every night.
SPEAKER B: Ryan did too, and got pretty fat.

15. SPEAKER A: (I wonder if it was) Andrew or/and Liam who always wore a kilt.
SPEAKER B: Duncan did too, thinking he had nice legs.

16. SPEAKER A: (I wonder if it was) Jeremy or/and Spencer who went to law school.
SPEAKER B: James did too, and earned top honors.

17. SPEAKER A: (I wonder if it was) Nicholas or/and Dan who played Angry Birds.
SPEAKER B: Aiden did too, and thought it was great fun.

18. SPEAKER A: (I wonder if it was) Claire or/and Liz who spoke up against the new regulations. SPEAKER B: Minta did too, but didn't convince anybody.

19. SPEAKER A: (I wonder if it was) Kara or/and Emma who shopped at Big Y. SPEAKER B: Katie did too, and bought a week's worth of groceries.

20. SPEAKER A: (I wonder if it was) Gary or/and Tom who fed the dogs. SPEAKER B: Ryan did too, and then took them for a walk.

Example sentences from filler experiments.

1. Speaker A: That family at the vet's office was really upset. Speaker B: Yeah, apparently the cat that the dog bit had to be put down.

2. Sophia emphasized the plan for the new addition was scrapped way too early.

3. The man wanted to eat a sweet snack. However, he ate a pretzel that was covered in salt when he was at the carnival.

4. SPEAKER A: I'm confused, did Jimmy give money to a business man? SPEAKER B: I think he gave money to a hobo when he was in Central Park.

References

- Beaver, D.; Clark, B. *Sense and sensitivity: How focus determines meaning*. Wiley Blackwell Publishers; West Sussex, UK: 2008.
- Beaver, D.; Zeevat, H. Accommodation. In: Ramchand, G.; Reiss, C., editors. *The Oxford Handbook of Linguistic Interfaces*. Oxford University Press; Oxford: 2007. p. 503-538.
- Bergen L, Grodner D. Speaker knowledge influences the comprehension of pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 2012; 38:1450–1460.
- Bock K, Mazzella JR. Intonational marking of given and new information: Some consequences for comprehension. *Memory & Cognition*. 1983; 11:64–76. [PubMed: 6855561]
- Breheny R, Ferguson HJ, Katsos N. Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition*. 2013; 126(3):423–440. doi: 10.1016/j.cognition.2012.11.012. [PubMed: 23291422]
- Clark, HH.; Haviland, SE. Comprehension and the given-new contract. In: Freedle, R., editor. *Discourse comprehension and production*. Ablex Publishers; Norwood, N.J.: 1977. p. 1-40.
- Clifton C Jr. Situational context affects definiteness preferences: Accommodation of presuppositions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2013; 39(2):487–501. doi: 10.1037/a0028975.
- Clifton C Jr. Frazier L. Discourse Integration Guided by the Question under Discussion. *Cognitive Psychology*. 2012; 65:352–379. [PubMed: 22683609]
- Frazier, Lyn. Processing accommodation: The big fish in a small pond. Workshop on accommodation (Invited talk), The Ohio State University; Oct 13. 2006
- Frazier L. Do null subject sentences (mis-)trigger pro - drop grammars? *Journal of Psycholinguistic Research*. 2015 in press.
- Garrod S, Sanford A. The mental representation of discourse in a focused memory system: Implications for the interpretation of anaphoric noun phrases. *Journal of Semantics*. 1982; 1:21–41.
- Ginzburg, J. *The Interactive Stance: Meaning for Conversation*. Oxford University Press; Oxford: 2012.
- Goodwin G, Johnson-Laird PN. Mental models and Boolean concepts. *Cognitive Psychology*. 2011; 63:34–59. [PubMed: 21679784]
- Grant M, Clifton C Jr. Frazier L. The role of Non-Actuality Implicatures in processing elided constituents. *Journal of Memory and Language*. 2012; 66:326–343. [PubMed: 22247589]
- Grodner, D.; Sedivy, JC. The effect of speaker-specific information on pragmatic inferences. In: Pearlmuter, N.; Gibson, E., editors. *The processing and acquisition of reference*. MIT Press; Cambridge, MA: 2015. in press
- Hamblin C. Questions in Montague English. *Foundations of Language*. 1973; 10:41–53.
- Haviland SE, Clark HH. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*. 1974; 13:512–521.
- Hawkins, JA.; Karttunen, L. Definiteness and indefiniteness: A study in reference and grammaticality prediction. *Linguistics and Philosophy*. Vol. 1. Croom Helm; London: 1978. 1977. Syntax and semantics of questions; p. 3-44.
- Karttunen L. Syntax and semantics of questions. *Linguistics and Philosophy*. 1977; 1:3–44.
- Lewis D. Scorekeeping in a language game. *Journal of Philosophical Logic*. 1979; 8:339–359.
- Mack J, Clifton C Jr. Frazier L, Taylor PV. Pragmatic constraints influence the restoration of optional subjects. *Journal of Memory and Language*. 2012; 67(1):211–223. [PubMed: 22754087]
- Piantadosi ST, Tily H, Gibson E. The communicative function of ambiguity in language. [Research Support, U.S. Gov't, Non-P.H.S.]. *Cognition*. 2012; 122(3):280–291. doi: 10.1016/j.cognition.2011.10.004. [PubMed: 22192697]

- R_Development_Core_Team. A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2012. <http://www.R-project.org>
- Rayner K, Sereno S, Morris R, Schmauder R, Clifton C Jr. Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*. 1989; 4(3-4):21–50.
- Rayner K, Sereno SC. Regression-contingent analyses: A reply to Altmann. *Memory & Cognition*. 1994; 22:291–292.
- Rayner K, Warren T, Juhasz BJ, Liversedge SP. The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 2004; 30:1290–1301.
- Roberts C. Information structure in discourse: Towards an integrated theory of pragmatics. *Semantics and Pragmatics*. 2012/1996; 6:1–69. (Originally published as The Ohio State University Working Papers in Linguistics, 1996, 49.).
- Rohde, D. Linger: A flexible program for language processing experiments. 2003. Retrieved from <http://tedlab.mit.edu/~dr/Linger/>
- Rohde H, Horton WS. Anticipatory looks reveal expectations about discourse relations. *Cognition*. 2014; 33:667–691. [PubMed: 25247235]
- Rooth, M. Association with focus. Unpublished Dissertation; Department of Linguistics, University of Massachusetts, Amherst, Ma: 1985.
- Rooth M. A Theory Of Focus Interpretation. *Natural Language Semantics*1. 1992:75–116.
- Schwarz, F., editor. Experimental perspectives on presuppositions. Springer International Publishing; New York: 2015a.
- Schwarz, F. Introduction: Presuppositions in context -- theoretical issues and experimental perspectives. In: Schwarz, F., editor. Experimental perspectives on presuppositions. Springer; Heidelberg and New York: 2015b. p. 1-38.
- Singh R, Fedorenko E, Mahowald K, Gibson E. Accommodating presuppositions is inappropriate in implausible contexts. *Cognitive Science*. 2015 in press.
- Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. Integration of visual and linguistic information in spoken language comprehension. *Science*. 1995; 268:1632–1634. [PubMed: 7777863]
- Tian Y, Breheny R, Ferguson HJ. Why we simulate negated information: A dynamic pragmatic account. *Quarterly Journal of Experimental Psychology*. 2010; 63(12):2305–2312.
- Van Berkum JJA, van den Brink D, Tesink CMJY, Kos M, Hagoort P. The neural integration of speaker and message. *Journal of Cognitive Neuroscience*. 2008; 20:580–591. [PubMed: 18052777]
- van Kuppevelt J. Inferring from topics: Scalar implicatures as topic dependent inferences. *Linguistics and Philosophy*. 1996; 19:393–443.
- Zondervan, A.; Meroni, L.; Gualmini, A. Experiments on the role of the question under discussion for ambiguity resolution and implicature computation in adults. In: Friedman, T.; Ito, S., editors. *Proceedings of SALT XVIII*. Cornell University Press; Ithica, NY: 2008. p. 765-777.

Highlights

- Disjunctive phrases are less natural and read more slowly than conjunctive phrases
- This penalty is reduced or eliminated in future tense sentences
- The penalty is reduced in a context indicating the author's uncertainty
- Disruption appears on early eyetracking measures of the disjunct
- It is difficult to accommodate an epistemic state to justify the disjunction

Table 1

Means (and standard errors) of naturalness/coherence ratings for Experiments 1 - 4 (7-point rating scale, 7 = natural/coherent)

	Items with 'or' in subject	Items with 'and' in subject	Difference
Experiment 1	4.46 (0.087)	5.80 (0.068)	1.34
Experiment 2			
No question	4.92 (0.10)	5.81 (0.07)	0.89
Explicit ques	5.48 (0.09)	6.21 (0.08)	0.73
Experiment 3			
Continuation	4.40 (0.11)	6.00 (0.08)	1.60
No continuation	5.27 (0.11)	6.49 (0.70)	1.22
Experiment 4A			
	Past tense		
Continuation	4.51 (0.16)	5.95 (0.11)	1.44
No continuation	4.54 (0.16)	6.34 (0.09)	1.80
	Future tense		
Continuation	5.02 (0.15)	6.05 (0.10)	1.03
No continuation	5.15 (0.14)	6.13 (0.11)	0.98
Experiment 4B			
	Past tense		
Continuation	4.75 (0.14)	6.13 (0.10)	1.38
No continuation	4.75 (0.13)	6.30 (0.09)	1.55
	Future tense		
Continuation	4.87 (0.13)	5.94 (0.10)	1.07
No continuation	5.00 (0.13)	6.13 (0.11)	1.13

Table 2

Eyetracking Measures, Mean Times (ms) or Proportions, with SEs, Experiment 5 ('cleft' indicates presence of *I wonder if it was...* lead-in)

Condition	Region 3 (X and/or Y)	Region 4 (Rest of S1)	Region 6 (Z did too...)	Region 7 (end of S2)
First Pass				
Or, Simple	562 (15)	946 (33)	887 (21)	402 (15)
Or, Cleft	416 (13)	1025 (34)	910 (24)	420 (20)
And, Simple	513 (14)	970 (35)	926 (19)	441 (18)
And, Cleft	424 (13)	946 (32)	898 (26)	426 (18)
Go Past				
Or, Simple	585 (17)	1190 (36)	1112 (35)	838 (48)
Or, Cleft	513 (17)	1176 (40)	1178 (41)	879 (50)
And, Simple	529 (14)	1155 (36)	1079 (32)	761 (39)
And, Cleft	546 (19)	1230 (37)	1207 (35)	805 (43)
Total Time				
Or, Simple	759 (14)	1176 (35)	1161 (28)	501 (21)
Or, Cleft	586 (18)	1185 (35)	1226 (32)	495 (22)
And, Simple	667 (22)	1136 (33)	1088 (22)	504 (20)
And, Cleft	620 (20)	1185 (31)	1210 (29)	488 (20)
Regressions Out				
Or, Simple	0.03 (0.01)	0.23 (0.02)	.141 (.020)	.450 (.032)
Or, Cleft	0.15 (0.02)	0.13 (0.02)	.156 (.022)	.511 (.034)
And, Simple	0.03 (0.01)	0.20 (0.02)	.124 (.020)	.372 (.032)
And, Cleft	0.18 (0.02)	0.23 (0.03)	.206 (.025)	.447 (.034)