Published in final edited form as:

Cogn Psychol. 2012 September ; 65(2): 352–379. doi:10.1016/j.cogpsych.2012.04.001.

Discourse Integration Guided by the ‘Question under Discussion’

Charles Clifton Jr. and Lyn Frazier

University of Massachusetts Amherst

Abstract

What makes a discourse coherent? One potential factor has been discussed in the linguistic literature in terms of a Question under Discussion (QUD). This approach claims that discourse proceeds by continually raising explicit or implicit questions, viewed as sets of alternatives, or competing descriptions of the world. If the interlocutor accepts the question, it becomes the QUD, a narrowed set of alternatives to be addressed (Roberts, 1996/To appear). Three eye movement recording studies are reported that investigated the effect of a preceding explicit QUD (Experiment 1) or implicit QUD (Experiments 2 and 3) on the processing of following text. Experiment 1 revealed an effect of whether the question queried alternative propositions or alternative entities. Reading times in the answer were faster when the answer it provided was of the same semantic type as was queried. Experiment 2 tested QUDs implied by the alternative description of reality introduced by a non-actuality implicature trigger such as *should X* or *want to X*. The results, when combined with the results of Experiment 3 (which ruled out a possible alternative interpretation) showed disrupted reading of a following verb phrase that failed to resolve the implicit QUD (*Did the discourse participant actually X?*), compared to reading the same material in the absence of a clear QUD. The findings support an online role for QUDs in guiding readers’ structuring and interpretation of discourse.

Keywords

Discourse comprehension; discourse coherence; eyetracking during reading; question-answer pairs; implicatures; Question under Discussion

We argue that discourse coherence is enhanced when its content is relevant to a ‘question under discussion’ (QUD) that has been explicitly or implicitly introduced. We present evidence that comprehension of the discourse is facilitated when following material addresses a currently-likely QUD, and describe some factors that make a particular QUD likely. The notion of QUD comes from an analysis that views discourse structure as a series of questions, implicit or explicit, which if accepted by the interlocutor are ideally followed by comments or replies that partially or completely answer them (Beaver & Clark, 2008; Roberts, 1996/To appear, 2004). An explicit question-answer pair, where the question is overt and the answer is a direct reply, is a straightforward example. But a QUD can be introduced implicitly or inferentially, or it could even be established after the fact, as by the client’s second sentence in (1) (adapted from Cooper, Engdahl, Larsson, & Ericsson, 2000), which relies on cultural knowledge of travel agents.

(1) Travel agent: Where will you be traveling?

Client: Madrid. There will be two of us.

Many factors, to be sure, contribute to discourse coherence. Some factors, including topichood, focus, order of mention, and various aspects of information status including the given vs. new contrast, affect the coherence of referring to particular entities or events (Gernsbacher, 1997; O'Brien & Myers, 1999; Sanford & Garrod, 1990, among others). Other factors, including causal coherence and temporal and spatial contiguity affect the ease of relating distinct propositions in a text to one another (McCrudden, Magliano, and Schraw, 2011; Trabasso & van den Broek, 1985). Useful overviews of this work appear in Garrod & Sanford, 1994 and Zwaan & Rapp, 2006). A QUD has a different status from these factors. It can be defined in largely formal terms, as a set of alternatives introduced by independently required syntactic and semantic processing of an utterance. As we will attempt to explain, a likely QUD can be identified in advance and can guide a reader or listener's expectations for upcoming semantic content. Meeting these expectations leads to increased discourse coherence.

Much of the development of the QUD notion (Beaver & Clark, 2008; Roberts, 1996/To appear, 2004) analyzes each step of a discourse as introducing a potentially very large family of possible QUDs, and later material as selecting from among these by a sort of 'backwards inference' (the *two of us* answer in (1) illustrates this). Such a perspective on QUDs has been very useful in analyzing how discourses can be interpreted as being coherent. However, viewing a QUD as a question selected only 'after the fact' from a large set of possible QUDs does not allow the notion to play a computationally tractable role in guiding a reader or listener's expectations about upcoming discourse material. In contrast, we emphasize a different perspective on QUDs. We suggest that readers may consider only a small number of likely QUDs, ones that are introduced by independently required syntactic and semantic computations. Rather than enumerating a large family of possible QUDs, this method will generate just a few QUDs, permitting them to constrain assumptions about following discourse.

We base this perspective on the currently widely accepted semantic analysis of questions, which claims that the denotation of a question is the set of possible (or true) answers. (Hamblin, 1973; Karttunen, 1977; Rooth, 1992). In such an analysis, the meaning of *Who left?* is something like { *Tom left, Dick left, Harry left...* }. The denotation of *Did Tom leave* is { *Tom left, Tom didn't leave* }. That is, the denotation of a question is a set of alternatives; a legitimate answer is the selection of one or more of these alternatives. A QUD that is introduced by an explicit or implicit question is simply the set of alternatives that potentially answer the question. An answer that selects an alternative or alternatives is expected, and contributes to discourse coherence. However, formal alternatives can be introduced by other syntactic and pragmatic devices such as contrastive focus and focus particles like *only*. Other alternatives relevant to our analysis of QUDs can be introduced by grammatical devices such as disjunction (e.g., *Tom or Dick left*) or indefinite reference (*I bet on a horse*, as opposed to *I bet on the horse*), and (as we illustrate below) certain implicatures.

The power of an analysis in terms of QUDs is that independently-motivated linguistic analyses propose that various linguistic structures and devices can introduce alternatives. A reader presented (perhaps implicitly) with a set of alternatives can legitimately expect the following discourse to select from among them. Meeting this expectation will increase the coherence of a discourse and, we assume, facilitate its comprehension. In the General Discussion, we briefly discuss some limitations on the power of a QUD to guide comprehension. Most saliently, a reader must be responsive to the QUD, which may depend on the structure of the discourse (e.g., material in the main assertion of a sentence is more likely to guide comprehension than material in a subordinate clause or 'not at issue')

material), and a reader must be sufficiently involved and motivated to attempt to find coherence in a discourse.

The expectations introduced by a QUD are not primarily lexical or syntactic expectations, whose influence on discourse comprehension has been demonstrated in various ways (Kamide, 2008). Rather, the expectations involve alternative semantic or pragmatic objects, e.g., the alternative answers to a question. Any QUD-based syntactic expectations are dependent upon the semantic alternatives offered by a QUD. For instance, if a QUD involves a set of alternative propositions, a reader may expect a clause; if the alternatives are entities, the reader may expect a noun phrase.

Some empirical evidence does exist that is directly interpretable in terms of a QUD. Zondervan, Meroni, and Gualmini (2008) reported a series of experiments explicitly designed to test the claim that the current QUD influences the interpretation of an ambiguous sentence or one that permits a scalar implicature.

Rohde, Levy, and Kehler (2011) provided experimental evidence that discourse coherence relationships such as Causal and Contiguity relationships play roles in the interpretation and processing of relative clauses, and Rohde & Kehler (2009) explicitly recognized the relation between discourse coherence relationships and QUDs

Grant, Clifton, & Frazier (2012) demonstrated that QUDs are not necessarily introduced by questions, but can also be created as inferences or implicatures. They proposed that a QUD could be introduced by a novel type of implicature (one kind of inference that can be cancelled) termed a 'non-actuality implicature' (NAI). Such an inference implicates a state of affairs that is an alternative to the state of affairs in the actual world, and such a set of alternatives constitutes a QUD. As will be described in the Introduction to Experiment 2, Grant et al. found that reading of subsequent elliptical material that addressed these alternatives was facilitated, compared to reading the same material following a sentence that did not introduce the alternatives.

As mentioned above, a QUD can be introduced in various ways. In the present work, we rely on the two ways we have been discussing: QUDs can be introduced by explicit questions, and they can be introduced as NAIs by assertions containing modals such as *could have* and *should have*. We present three experimental demonstrations that the fluency with which a discourse is read is affected by whether or not it appropriately addresses the QUD that the discourse had established. Our first experiment establishes a QUD with an explicit question, a question whose meaning is strongly biased either to involve a set of alternative entities, or to involve a set of alternative propositions. A following answer then mentions either entities or a proposition. We expect that answers that match the semantic content of a question will be read with less disruption than answers that mismatch. Our second and third experiments establish a QUD via a NAI (non-actuality implicature), and examine the reading of a short discourse where later material first apparently answers the QUD but then this apparent answer is shown not to be the actual content of the discourse. Late disruption is expected in this case, compared to a condition in which the QUD was not established initially. All experiments are based on the non-controversial assumption that textual material that is relevant to preceding material will be understood more fluently than material that is not relevant. Their intended contribution is to explore the proposal that introducing a set of alternatives into the discourse introduces a QUD, and addressing this QUD is an important basis of relevance and therefore discourse coherence.

Experiment 1

The first experiment relies on a large literature that demonstrates that subcategorization preferences, e.g., whether a particular verb occurs more often with a direct object or with a sentence complement, influences sentence processing quickly (Clifton, Frazier & Connine, 1984; Ferreira & Henderson, 1990; Ford, Bresnan, & Kaplan, 1983; Garnsey, Pearlmutter, Myers, & Lotocki, 1997; Kennison, 2001; Trueswell, Tanenhaus, & Kello, 1993). Sentences where the subcategorization preference of a lexical item matches the form of its complement are read more fluently than sentences with mismatches. The results are often interpreted in terms of a processing system where the preferred subcategorization of a verb immediately influences the syntactic structure assigned to the material following the verb, but some evidence suggests that the effect may reflect differential plausibility of items (Kennison, 2001) or that the subcategorization preferences may primarily exert an influence on the evaluation of structure already assigned (Frazier, 1989).

Our experiment is not designed to address these differences in interpretation, but simply relies on the empirical fact that subcategorization preference matches affect reading time. We investigate whether this matching effect is modulated by the subcategorization preference of the verb in a preceding question. We predict that it should be: if the subcategorization preference of the question verb affects the QUD that the question establishes, and if the answer addresses this QUD, reading should be more fluent than if the answer does not address the presumed QUD.

Concretely, we had subjects read question-answer discourses like those in (2) (see Table 1 for a full paradigm, containing different examples). In this example, Speaker A asks a question with a verb (*say*) that is biased toward a sentence complement (e.g., *Joe said her name was Marie*) rather than a noun phrase complement (*Joe said her name*). We suggest that this bias makes the subject inclined to interpret the question as asking about a set of alternatives that differ in the propositions they involve (statements describing states or events, or more technically, statements that can be true or false), not only with respect to the alternative entities they involve (possible referents of noun phrases). Speaker B then answers the question with either a sentence that syntactically provides an answer that is (a) the statement of a proposition (*the rumor should have been stopped sooner*) or (b) a noun phrase naming an entity (*the rumor*). The former matches the presumed QUD of the preceding question in (2), and should be read with less disruption than the latter.

(2) SPEAKER A: What did the CIA director say?

SPEAKER B: (a) He confirmed the rumor should have been stopped sooner.

(b) He confirmed the rumor once it had spread widely.

The experiment contrasted items like (2) with items that were minimally different, differing only in the subcategorization bias of the verb in the question. For instance, the direct object-biased verb *write* could be substituted for the verb *say* in (2). This would presumably establish a QUD for an entity, not a proposition. The experiment also varied the bias of the verb in the answer (*confirm* in the example; see Table 1). A common finding in the subcategorization preference literature cited above is that a sentence is read faster when the subcategorization preference of its verb matches the sentence's syntactic form. Our experiment was not designed to provide a clear test of this, since we did not attempt to control lexical material in the conditions being contrasted by (e.g.) crossing verbs and their complements. The experiment was instead designed to provide a clean test of whether the QUD introduced by an explicit question affected the reading time of a following sentence

that did or did not address the QUD, and a preliminary look at whether any such effect interacted with the bias of the verb in the answer.

Method

Materials—40 sets of question-answer pairs were constructed, as illustrated in Table 1. All materials were derived from those used in Gahl and Garnsey's (2004) study of the effects of syntactic construction frequency on pronunciation. Gahl and Garnsey provided norms of the frequency with which each of a set of verbs was used with an NP (direct object noun phrase) complement and with a finite-S (sentence) complement, and constructed sentences that used each verb but continued with either a direct object or a sentence complement. We based the sentences we used as answers to questions on the Gahl and Garnsey sentences, modifying them to make them suitable for eyetracking purposes (e.g., it was necessary to have several words following the critical word so that it did not end the sentence) and to allow them to serve as answers to a preceding question. Our questions were constructed using verbs from the Gahl and Garnsey materials, but contained new content otherwise. All materials that we used appear in Appendix 1.

Thirteen verbs that had higher frequencies of direct object than sentence complements in the Gahl and Garnsey norms (means of 69 vs. 7%) were chosen, together with 18 verbs with higher sentence complement frequencies (57 vs. 16%). Because of the limited number of available verbs and the need to maintain plausibility, most were used in both questions and answers, and some were used in up to seven different questions and three different answers. No verb appeared in both members of a single question-answer pair. There were 20 items with direct object biased verbs in the answer (as illustrated in the top half of Table 1), and 20 with sentence complement biased verbs in the answer (bottom half of Table 1). Each answer had two syntactic versions, one with a direct object syntactic resolution, and one with a sentence complement resolution. There were two questions for each item, differing only in that one contained a verb that was biased toward a direct object interpretation, and one a verb that was biased toward a sentence complement.

Since there were 20 direct object biased answer sets, and 20 sentence complement biased sets, each of the 40 items occurred in four different forms defined by the bias of the preceding question crossed with the syntactic resolution of the answer. The resulting items were combined with a total of eight two-sentence practice items and 56 two-sentence discourses from other unrelated experiments.

Subjects—Sixty University of Massachusetts undergraduates signed written informed consent forms and were tested (receiving course credit) in individual 45-minute sessions during which their eye movements were recorded while reading the question-answer pairs. Four subjects were eliminated for having over 40% errors, two for having too many track losses, three for not following instructions (e.g., not reading the question of a question-answer pair), and four for being unable to track. Because one subject had to be rejected after the experiment was completed, 47 subjects remained, with equal numbers assigned to each of the four counterbalancing conditions to be described with one exception.

Design and Procedures—Four counterbalanced lists were prepared, so that each of the 20 items with direct object biased answers appeared in the one of the four conditions in each list, and similarly for the 20 items with sentence complement biased answers. The design had question-verb bias and syntactic resolution of the answer as crossed within-item factors, and answer-verb bias as a between-item factor. All factors were within-subjects. Each list thus contained five items in each of the eight conditions of the experiment. Since different sentences with different words and phrases had to be used in the different answer-verb and

syntactic resolution conditions, the experiment was not designed to compare these directly. Rather, the experiment was designed to explore effects and interactions involving question bias.

The 40 items that a subject saw appeared in individually-randomized order, randomly interleaved with items from the other experiments. The question and answer appeared on separate lines of a display monitor. The question was preceded by “SPEAKER A:” and the answer by “SPEAKER B:”. Subjects read the question and the answer for understanding. Five question-answer pairs from each group of 20 were followed by a simple two-choice question to ensure comprehension.

Subjects’ eye movements were recorded while reading the sentence. They were seated 55 cm from a CRT monitor that displayed the sentences. Their eye movements were recorded at 1 kHz using an EyeLink 1000 (SR Research, Toronto, Ontario, Canada) eyetracker controlled by a PC running the University of Massachusetts EyeTrack software (<http://www.psych.umass.edu/eyelab>). All items were presented in two lines, in a fixed-width font (Monaco). When they had finished reading an item, subjects were instructed to look to a colored square affixed to the right edge of the monitor and to pull a trigger on a game controller. A two-choice comprehension question then appeared on the screen, which subjects answered by pulling the left or right trigger on a game controller to indicate whether the left or the right answer was correct. Calibration was checked on each trial by checking the alignment with a single centered target. Minor errors were automatically corrected; unacceptably large errors led to recalibration. The entire experiment took 45 minutes or less.

Results

Overall accuracy on the comprehension questions averaged 80%, indicating adequate comprehension. The eye movement data were cleaned of track losses, blinks, and long fixations over 800 ms using the program EyeDoctor (<http://www.psych.umass.edu/eyelab>), and fixations less than 80 ms were deleted unless they could be incorporated into an adjacent fixation within 3 characters (see Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). Finally, trials with blinks or track losses on the critical word were eliminated, 9.2% of all trials.

The answer sentences were divided into several regions for analysis purposes: The main clause subject plus its verb (region 1), the noun phrase (DP, or ‘determiner phrase’) following the verb (Region 2), the disambiguating region (the first word or series of words following the DP that contained at least 8 letters; Region 3); and the remainder of the sentence (sometimes null; Region 4). The three most informative eyetracking measures are presented in Table 2 (see Rayner et al., 1989, for descriptions of these measures). Region 3 was the primary target of statistical analysis, because it was the only region present in all items that could be affected in theoretically interesting ways by the question bias.

Each measure was analyzed as a linear mixed model, using the program lmer (Bates, 2005) under the R programming language (R Development Core Team, 2007). The fixed effect factors were verb bias in the question (QBias), verb bias in the answer (ABias), and syntactic resolution (Syntax). Sum coding (Crawley, 2007; also termed ‘contrast’ or ‘ANOVA-style’ coding) was used, so that the effect of each factor was evaluated at the mean of the other factors. The random factors were subjects and items. Models with random intercepts were initially compared with models that had random intercepts plus random slopes involving the interaction of all three factors. If the more complex model was significantly better than the simple model by a likelihood ratio test (Baayen, 2008), it was simplified by progressively removing interaction terms and eventually main effect terms from the random factors as long as the more complex model was not significantly better than

the simplified model. The resulting random factor model will be specified in each analysis. Since mcmc approximations to p values are not available for models with random slopes, values of $t > 2.0$ will be treated as significant (since the t distribution approximates the z distribution with the large df 's in our models). z and p values however are reported for analyses of the proportions of regressions out, where logit analyses were used (Jaeger, 2008).

Go-past time—The most informative data were provided by the go-past measure, which accumulates first pass time and any time spent regressing from the region being analyzed and re-reading that region before progressing in the text. The best model for the analysis of the disambiguating region (Region 3) had all three fixed effect factors interacting in random slopes for subjects, and with just the syntactic resolution factor ("Syntax") contributing to random slopes for items. Our critical predictions involved effects and interactions involving QBias, the subcategorization bias of the question. The interaction between QBias and Syntax was significant ($t = 2.62$), and is illustrated in Figure 1, which presents the Syntax factor as a match (M) or mismatch (MM) with the QBias factor. Go-past times were significantly lower when the syntactic resolution of a sentence matched the question bias than when it did not.

There was also an interaction between ABias and Syntax ($t = 3.49$), reflecting the previously-observed effect of verb bias or of plausibility on sentence interpretation (Garnsey et al., 1997; Kennison, 2001). While resolution as a direct object was faster than resolution as a sentence complement when the answer verb was biased toward a direct object (551 vs 700 ms), the opposite held true when the answer verb was biased toward a sentence complement (855 vs 629 ms; see Table 2 for a detailed presentation of the data). We do not explore this finding further, because the experiment was not designed to allow a comparison of conditions that contained the identical lexical material.

Analyses of the postdisambiguating region were uninformative. The best model was the same as the best model for the disambiguating region. There were significant effects of ABias ($t = 2.56$) and Syntax ($t = 2.17$) (sentences with direct object biased verbs or with direct object resolution were read more slowly than ones with sentence complement bias or resolution). However, these differences are difficult to interpret because of uncontrolled differences between the materials in the different conditions.

First pass time—The observed effect on go-past time can be partially decomposed into effects on first pass time and on the probability of regressing out of a region. The analysis of first pass times in the disambiguating region used a model with random intercepts for subjects, and with just the syntactic resolution factor contributing to random slopes for items. The QBias x Syntax interaction that was present in go-past times was nonsignificant in first pass times ($t < 1.0$). Items with sentence complement biased answer verbs were read more slowly than items with direct object biased answer verbs (484 vs. 419 ms; $t = 2.31$), but this effect was qualified by a significant interaction of ABias and Syntax and a significant interaction among all three factors ($t = 2.47$ and $t = 2.53$, respectively). The former interaction was due to long times for answers with sentence complement biased verbs that were resolved as having simple direct objects (524 ms vs. means of 402 to 443 ms for the other conditions). Examination of Table 2 suggests that this interaction was larger when the question had a direct object verb bias than when it had a sentence complement verb bias. However, both of these effects are difficult to interpret because of the different materials in the different conditions, but a good part of the effects may be due to the relative implausibility of the answers with sentence complement biased verbs that were resolved as direct object sentences.

The only first pass time effect observed in the final region (analyzed with random subject intercepts and interacting random item intercepts and slopes) was the effect of verb bias in the answer. Sentences that had direct object biased verbs were read more slowly than those with sentence complement biases, 529 vs. 428 ms, $t = 2.60$. Once again, the difference might be due to uncontrolled item differences.

Regressions out—While first pass times apparently did not contribute to the QBias x Syntax interaction observed in go-past times, the frequency of first-pass regressions from the disambiguating region did. The QBias x Syntax interaction was significant ($t = 2.80$, $p < .01$), and is illustrated in Figure 2. It has the same form as Figure 1: Regressions out were more frequent when the syntactic resolution of the answer matched the bias of the question, and thus satisfied the presumed QUD.

In addition, the interaction of ABias and Syntax was significant ($t = 3.79$, $p < .001$). Regressions were markedly more frequent when the syntactic resolution mismatched the bias of the answer verb than when it matched, again presumably reflecting differences in plausibility of the items. A similar effect ($t = 3.26$, $p < .001$) together with a main effect of ABias ($t = 2.35$, $p < .02$) was present in the (very frequent) regressions out of the final region of the sentence.

Discussion

The primary result of the experiment indicated the role of a QUD in facilitating interpretation: reading of the disambiguating region was more fluent when the syntactic disambiguation was consistent with the bias of the verb in the question than when it was inconsistent. Inconsistency resulted in a higher frequency of regressions out of the disambiguating region and a longer time from first entering that region to going past it. We interpret this finding as showing that a direct-object biased verb in a question sets up a QUD that involves a set of alternative entities, and a sentence-complement biased verb in a question sets up a QUD that involves a set of alternative propositions. When the answer sentence is disambiguated toward a direct object interpretation, the former is satisfied; when it is disambiguated toward a sentence complement interpretation, the latter is satisfied.

The other results in the experiment lack clear interpretations, because the experimental materials were not designed to investigate the factors involved. However, there were consistent interactions between answer verb bias and answer syntactic resolution, indicating that reading was disrupted when syntactic resolution mismatched the bias of the verb in the answer and, in first pass times, an interaction that also involved the bias of the verb in the question (a three-way interaction that disappeared in go-past time). The two-way interaction could be taken as evidence that subcategorization bias guided analysis of the sentence or that implausibility affected reading time, or both (Garnsey et al., 1997; Kennison, 2001; Trueswell et al., 1993).

We have discussed our results in terms of semantic expectations: a QUD about a proposition helps a reader integrate an answer that provides a proposition, etc. It is possible that the expectations are, instead, syntactic. It is even possible that the results could be interpreted in terms of syntactic priming in comprehension (Branigan, Pickering, Liversedge, Stewart, & Urbach, 1995; Branigan, Pickering, & McLean, 2005; Frazier, Taft, Roeper, Clifton, & Ehrlich, 1984). This could, in principle, prime the assignment of a similar subcategorization to the verb in the answer, but we think that this account is unlikely for several reasons, including the fact that syntactic priming in comprehension is difficult or impossible to obtain when the prime and target verbs differ as they did in our materials.

Our main conclusion, then, is that while a variety of factors certainly affect the time to read discourses (and have been explored in other research), we have demonstrated the importance of a QUD that is introduced by an explicit question, which was affected by the syntactic bias of the verb used in the question. When the verb of the question is most commonly used with a nominal direct object, the QUD involves the identity of an entity, and is best satisfied by an answer whose focus is a nominal phrase. When the verb is most commonly used with a sentence complement, the QUD involves a proposition, and is satisfied by an answer whose focus is a sentential complement.

Experiment 2

In ordinary discourse, most QUDs are presumably introduced indirectly, not by direct questions. They can be introduced in many ways, including statements introducing reasons or causes or statements indicating that a succession of events will be described (Kehler, 2002), statements introducing disjunctions, indefinite noun phrases introducing a new and incompletely identified entity into the discourse, statements marked by contrastive focus, etc. We (Grant et al, 2012) have proposed that they can also be introduced by phrases containing modals (e.g., *should have*, *could have*, *must be*) as well as some verbs and adjectives (e.g., *wanted to*, *is eager to*). Such phrases carry the implicature that the event of their complement (e.g., *go to Paris*) hasn't actually taken place. They therefore set up the alternatives that the event did or did not take place, which, we have suggested, introduces a QUD into the discourse: did the event (or state of affairs) described actually take place?

Grant et al. examined the reading of short discourses ending in an elliptical clause that mismatched its antecedent syntactically (e.g., *The material should have been released, but Gorbachev didn't*) and are arguably ungrammatical (Arregui, Clifton, Moulton, & Frazier, 2006; cf. Johnson, 2001, for an overview). Sentences with such clauses were found to be less acceptable than sentences with a syntactically parallel antecedent of the ellipsis, but the decline in acceptability (and the increase in reading time) was less when the antecedent contained a modal or modal-like phrase like those discussed above compared to when it contained only a simple assertion (e.g., *The material was released*). Grant et al. termed the implicature that the event in question had not actually occurred a 'non-actuality implicature' (NAI), and suggested first, that the NAI was computed on-line, and second, that it introduced a QUD which motivated the content of the elliptical clause. The elliptical clause directly addressed the QUD (whether or not the event in question took place), and therefore could easily be integrated into the discourse, partially alleviating the penalty introduced by the lack of a syntactic match between ellipsis and antecedent.

The present experiment more directly addresses whether a NAI can introduce a QUD that affects the interpretation of a following clause. It measured eye movements while people read simple 2-clause sentences like that in (3):

- (3) a. Jessica could go to Paris, but she hasn't, according to her sister, been to Rome any time recently.
- b. Jessica went to Paris, but she hasn't, according to her sister, been to Rome any time recently.

The first clause of sentence (3a) introduces a NAI (namely, that Jessica hasn't actually gone to Paris), which by our analysis, introduces a QUD (did she, or didn't she?). The next phrase *but she hasn't* seems to address this QUD, elliptically asserting that she did not in fact go to Paris. However, following the parenthetical *according to her sister*, the phrase *been to Rome* appears, and has to be interpreted as the explicit complement to *she hasn't*, thus displacing the elliptical interpretation under which that phrase addressed the QUD. Given our

assumption that readers attempt to integrate text in a way that answers a QUD, this should require a re-analysis toward the end of the sentence, appearing as disrupted reading late in (3a).

Sentences like (3b) served as controls. They did not introduce a NAI in the initial clause, and thus did not introduce a QUD that could be addressed by an elliptical *but she hasn't* (although their initial clause could be interpreted as introducing any of several other QUDs, depending on the discourse context: most obviously, who else did/didn't go to Paris, or where else did Jessica go). Because the phrase *but she hasn't* does not play an obvious role in the discourse (and if interpreted as an elliptical phrase, would contradict the immediately preceding assertion), its reading could be disrupted, but the material in the remainder of the sentence, after the parenthetical, resolves any problems that the reader may have encountered. This material indicates, after the fact, that the QUD that could integrate the discourse was something like 'where else has/hasn't Jessica gone.' This is not a QUD that could be anticipated (like the QUD introduced by an NAI or by an explicit question), but is one that, by backwards inference, helps to integrate the discourse. Thus, we predict that the later phrases of sentences like (3b) will be read faster than those of sentences like (3a).

Method

Materials—Twenty short one-sentence discourses were constructed in two versions each, as illustrated in Table 3. In one version (NAI), the discourse began with an initial clause (Region 1) that included the modal *could*, which supported a NAI. In the other version (non-NAI), the initial clause was in the simple past tense, not supporting an NAI. The discourse continued with a 'denial phrase' (Region 2) of the form *but s/he hasn't*, which if interpreted as an elliptical phrase could be taken to respond to the QUD presumably introduced by the NAI in the NAI condition, but (taken on its own) was superficially inconsistent with the initial simple past tense phrase. This was followed by a parenthetical like *according to what I hear* (Region 3), followed by a critical phrase like *visited her parents* (Region 4). This phrase was designed to contain a complement that contrasted with the complement of the verb phrase of the clause that introduced the discourse. It 'rescued' the apparent inconsistency between the first two phrases of the non-NAI condition, but contradicted the interpretation of Region 2 that could have been taken to address the QUD in the NAI condition. The sentence continued with a final phrase (Region 5).

These 20 discourses were combined with 8 practice discourses and 106 other filler discourses and discourses from other experiments. Four practice discourses and 33 other discourses were followed by simple 2-choice questions to ensure comprehension.

Subjects—Fifty-five University of Massachusetts undergraduates signed written informed consent forms and were tested in individual 40-minute sessions. Seven subjects were eliminated for having many track losses on over 25% of the critical items, and six for being unable to track. Approximately equal numbers of the remaining 42 subjects were assigned to each of the two counterbalancing conditions to be described.

Design and Procedures—Two counterbalanced lists were prepared, so that each of the 20 items appeared in one version in one list. The design had a single within-item factor, NAI vs. non-NAI introductory clause. Each list thus contained 10 items in each of the two conditions of the experiment. The 20 items that a subject saw appeared in individually-randomized order, randomly interleaved with items from the other experiments. The initial clause appeared on the first line of the display monitor, with the remainder (from Region 2 on) on the second line. Subjects read the question and the answer for understanding. Five of

the experimental discourses were followed by a simple two-choice question to ensure comprehension.

Subjects' eye movements were recorded while reading the sentence, following the procedures described for Experiment 1.

Results

Overall accuracy on the comprehension questions averaged 97%, indicating good comprehension. The eye movement data were cleaned of track losses, blinks, and long fixations over 800ms as described for Experiment 1. Finally, trials with blinks or track losses on the critical region, Region 4, were eliminated, 6.7% of all trials.

The discourses were divided into several regions for analysis purposes, as illustrated in Table 3. The most informative eyetracking measures are presented in Table 4. Different regions were subjected to statistical analysis of the different measures, as appropriate. Model comparison and selection was conducted as described for Experiment 1.

The overall pattern in the data was clear. In Region 3, the non-NAI condition took longer to read, in several measures, presumably indicating the clash between the first sentence and the polarity of the second/current sentence. But in later regions, the effects flipped, with the NAI condition taking longer, presumably because disambiguating material forced the processor to abandon the incorrect ellipsis analysis of *hasn't*, which had been encouraged by the NAI trigger.

First pass and total time—These measures were analyzed with Condition (NAI vs non-NAI) and Region (Regions 3, 4, 5) as factors. The levels of the Region factor were contrasted using Sum (contrast, or ANOVA-style) coding, which results in the effect of Condition being evaluated at the mean over all regions. No effects (apart from differences between regions, which reflect length and other lexical differences) approached significance in first pass time. The analysis of total time (using random intercepts and condition slopes of subjects and items), however, yielded a significant effect of condition ($t = 2.49$), which reflected the overall longer time for the NAI than for the non-NAI condition (means of 648 vs. 608 ms). This longer time appeared to show up only on Regions 4 and 5, but the interaction (tested using treatment coding rather than sum coding) between Condition and Region 3 vs. Region 4 only approached significance ($t = 1.89$) while the interaction between Condition and Region 3 vs. Region 5 was nonsignificant ($t = 1.01$). Despite the uncertain delay, it is clear that reading of material that contradicted an earlier apparent response to the presumed QUD was disrupted, as indicated by total times.

Second pass time—The significance of the total time effect, combined with the nonsignificance of the first pass time effect, points to an effect of Condition on re-reading time. Second pass, or re-reading, time deviates too far from a normal distribution (with a modal time of zero) to be analyzed using a linear mixed model, so average second pass times by subjects items and by items were computed and conventional F1/F2 analyses were performed, now on Regions 2, 3, and 4. The interaction between Condition and Region was significant ($F(2,82) = 10.97$, $p < .001$; $F(2,38) = 5.06$, $p < .002$) (but the main effect of Condition was nonsignificant, $p > .29$, and the main effect of Region is not of interest). Separate tests indicated that second pass time was longer for non-NAI than for NAI items on Region 2 ($F(1,41) = 4.31$, $p < .05$; $F(1,19) = 5.68$, $p < .03$), but shorter for non-NAI than NAI items on Region 3 ($F(1,41) = 6.18$, $p < .02$; $F(1,18) = 4.44$, $p < .05$) and marginally shorter on for non-NAI than NAI items on Region 4 ($F(1,41) = 3.43$, $p = .07$; $F(1,19) = 2.29$, $p < .15$).

Regressions out and go-past time—A logistic regression model (Jaeger, 2008) was used to analyze the proportions of regressions out of a region (with Region 3, 4, and 5 as a sum-coded variable and with random subject and item slopes and intercepts). While the main effect of Condition was nonsignificant ($z = 1.34$, $p > .18$), its interaction with the first contrast of Region (in which Region 3 is contrasted with the mean over all regions) was significant ($z = 3.04$, $p < .003$). In the non-NAI condition, regressions were disproportionately frequent from Region 3 compared to the other regions, resulting in the long second pass times for Region 2. This presumably reflects the clash between the apparent content of the initial clause and the denial in Region 2.

While regressions from the later regions were not particularly frequent for the NAI condition, go-past times (which reflect both the frequency of regressions out of a region plus the time spent reading the region and re-reading it and earlier regions) did appear to increase. This is best seen by coding the Regions factor as a numerical variable and evaluating its slope (in an analysis with random subject and item condition slopes). In such an analysis, the effect of Condition was significant ($t = 2.51$), indicating that go-past times were longer for the non-NAI than the NAI condition at the Regions zero-intercept, and the effect of Region slope ($t = 9.28$) and its interaction with Condition ($t = 3.25$) were both significant. Separate analyses done at each region indicated significantly longer go-past times for non-NAI than NAI in Region 3 ($t = 2.01$) and marginally significantly longer times for NAI than non-NAT in Region 4 ($t = 1.73$) and Region 5 ($t = 1.72$). Part of these arguably longer go-past times in the later regions presumably was spent re-reading material in Regions 3 and 4, resulting in the longer second pass times in these regions.

Discussion

Two opposite effects appeared in Experiment 2, one early and one late in the sentence. Reading of Region 3, the region immediately following the denial region, was disrupted in the non-NAI condition compared to the NAI condition, as reflected in frequent regressions from this region, long go-past times in it, and long second pass times in Region 2, the denial region. This presumably reflects a relatively quick response to the apparent contradiction between the initial clause (*Mary visited me*) and the next one (*but she hasn't*) interpreted as an ellipsis. However, this disruption did not persist. The later regions (and second pass reading of Region 3) showed a cost to the NAI condition as compared to the non-NAI condition. These regions introduced material that rationalized the apparent conflict between the first two clauses in the non-NAI condition, but that was inconsistent with an interpretation of Region 2 as confirming the presumed implicature of the first clause in the NAI condition, namely (in *Mary could visit me*) that Mary did not actually visit me.

This latter finding, we submit, provides evidence first, that an implicature (the NAI) is computed from the first modal clause, and second, that this implicature forms the basis of a QUD, namely, in the example that we have been using, whether or not Mary visited the speaker. When later material does not address this QUD (asserting, instead, that Mary visited someone else), it is integrated only with some difficulty into the ongoing discourse, requiring a re-interpretation of the denial phrase. This disruption does not appear immediately, but appears largely as long re-reading times initiated from the material that is inconsistent with interpreting the denial phrase as addressing the QUD.

It is possible, however, to take a somewhat different perspective on the results of Experiment 2, viewing the observed effect more as facilitation in the non-NAI condition than as integration difficulty in the NAI condition. If the denial phrase *but she hasn't* (see Table 3) is consistently interpreted as being an elliptical clause with the preceding clause (*Mary could visit me/visited me*) as its antecedent in both the NAI and the non-NAI conditions, it will successfully answer the QUD in the NAI condition, discharging the QUD,

but contradict the initial clause in the non-NAI condition. This latter case could plausibly introduce a new QUD (e.g., what else has she done? or what did she do?), which is addressed by the material in Region 4 of the non-NAI condition. This would predict faster reading times for Region 4 in non-NAI than in NAI, as observed. This interpretation is subtly different from the one we offer. Our proffered interpretation says that reading of the later material is slow in the NAI condition because it fails to address the QUD; the alternative says that reading of this material is fast in the non-NAI condition because it addresses a QUD that was raised by the earlier denial phrase. We address this difference in Experiment 3 by using materials that discourage interpretation of the denial phrase as being elliptical.

Experiment 3

The materials used in Experiment 3 were similar to those used in Experiment 2, apart from having a passive initial clause and introducing the NAI with a variety of different modals (*should*, *was supposed to*, etc.). They were adapted from materials used in Grant et al.'s. (2012) Experiment 1. Table 5 presents an example, with both NAI and non-NAI versions. The key difference from Experiment 2 is that the initial clause was in the passive voice, while the denial was active. This mismatch discourages treating the denial as elliptical (such an elliptical interpretation can be argued to be ungrammatical; Arregui et al., 2006). The alternative interpretation offered for Experiment 2 assumed that the denial was treated as elliptical, an interpretation that could be rejected in the non-NAI condition because of its inconsistency with the first clause. If the denial is not treated as elliptical (which we assume will generally be the case in Experiment 3), such rejection will not occur. Further, the Experiment 3 materials were such that even if the voice mismatch were overlooked and the denial treated as elliptical, it did not contradict the initial phrase, and thus would not introduce a new QUD in the non-NAI condition that would be addressed by the following material.

Our proffered hypothesis still predicts that reading time in the later region of the sentences will be disrupted in the NAI condition, compared to the non-NAI condition, because the material in this region does not address the QUD introduced in the NAI condition. The alternative hypothesis, by itself, does not predict a difference in reading time between the conditions. If the results prove to support our favored hypothesis, this would not mean that the alternative hypothesis played no role in Experiment 2, but would simply indicate that the processes that our favored hypothesis assumes play a role in Experiment 3, and presumably account for at least part of the effect in Experiment 2.

Method

Materials—20 pairs of items, as illustrated in Table 5, were constructed (all appear in the Appendix; one item, Item 16, had to be eliminated because of a typographical error). Each began with a passive clause, either a simple clause or one with a non-actuality implicature inducing modal or modal phrase. As in Experiment 2, this phrase was followed by an active denial phrase (e.g., *but Gorbachev didn't*) which could be elliptical except for its failure to match the preceding phrase in voice. This was followed by a parenthetical, set off by commas and spanning the two lines of the display, and then by a verb phrase that completed the second denial phrase in a way that did not address the QUD introduced by the initial clause (e.g., *notify the appropriate bureau*). A second sentence followed this first, so that the critical region (the verb phrase just mentioned) did not end the displayed text.

These items were combined with 106 items from other experiments or filler items (96 of which contained a display change, i.e., a region of text contained random letters until the reader's eye landed on it, at which time it reverted to the actual words of the sentence; see

Rayner et al., 1989). One quarter of the items were followed by simple 2-choice questions to ensure that subjects were reading for comprehension.

Subjects—Data are reported for 27 University of Massachusetts undergraduates who participated for course extra credit. An additional 2 subjects were rejected for making errors on over 40% of the questions, 16 for being difficult to track (because of the existence of display change items in the experiment, it was necessary to insist on being able to achieve a very good calibration), 2 for not being native speakers of English, and 4 for losing more than 40% of their trials because of blinks and track losses in the critical region.

Design and procedures—Two counterbalanced lists were prepared, so that each of the 20 items appeared in one version in one list. The design had a single within-item factor, NAI vs non-NAI introductory clause. Each list thus contained 10 items in each of the two conditions of the experiment. The 20 items that a subject saw appeared in individually-randomized order, randomly interleaved with items from the other experiments. The initial clause together with part of the parenthetical display appeared on the first line of the display monitor, with the remainder on the second line. Subjects read the question and the answer while their eye movements were tracked, as in the previous experiments.. Five of the experimental discourses were followed by a simple two-choice question to ensure comprehension.

Results

Overall accuracy on the comprehension questions averaged 86%, indicating adequate comprehension. The eye movement data were cleaned as described for Experiment 1. Finally, trials with blinks or track losses on the critical region, Region 4, were eliminated, 8.3% of all trials (considering only subjects whose data were retained for analysis).

The discourses were divided into several regions for analysis purposes, as illustrated in Table 5: the introductory clause through the initial comma (region 1), the denial (Region 2), the parenthetical region (Region 3), the critical region that addressed the content of the introductory clause (Region 4), and the remainder of the sentence (Region 5). The most informative eyetracking measures are presented in Table 6. Different regions were subjected to statistical analysis of the different measures, as appropriate. Model comparisons were conducted as described for Experiment 1, and the simplest model that was not significantly worse than a more complicated model is reported.

Several measures indicated that reading was disrupted in Regions 4 and 5 in the NAI conditions as compared to the non-NAI conditions. However, the significant disruption observed following the denial region (Region 2) of Experiment 2 was not observed in Experiment 3, eliminating the possibility that the denial region of the NAI condition of Experiment 2 satisfied the QUD of the initial clause or that the denial region of the non-NAI condition of that experiment introduced a new QUD that was satisfied by later material.

First pass, total time, and second pass time—In Regions 2 and 3, the effect of Condition approached significance for none of these measures. However, marginally significant differences appeared later. These measures were analyzed with Condition (NAI vs. non-NAI) and Region (Regions 4 and 5) as noninteracting factors. No effects (apart from differences between regions, which reflect length and lexical differences) approached significance in first pass time. The analysis of total time (using random condition slopes of subjects and items), however, yielded a marginally significant effect of Condition ($t = 1.854$) in an analysis with random condition slopes. This reflected the overall longer time for the NAI than for the non-NAI condition (means of 1122 vs. 1057 ms).

As in Experiment 2, second pass time was analyzed using conventional F1/F2 ANOVAs. No effects were convincingly significant in any region for second pass time. The closest approach to significance occurred in Region 4, where the NAI sentences were re-read longer than the non-NAI sentences by subjects ($F(1,26) = 4.35, p < .05$) but not by items ($F(2,18) = 1.27, p > .25$).

Go-past time and regressions out—Analysis of go-past time did provide clear evidence of disrupted reading in the NAI conditions. As was the case for first pass and total time, the effect of Condition was thoroughly nonsignificant in Region 2, the denial region ($t < 1.0$) and was nonsignificant in Region 3, the parenthetical region ($t = 1.27$). Thus, the apparent disruption for the non-NAI condition in this region cannot be trusted. However, a significant difference favoring the non-NAI condition appeared later. Treating Condition and Region as non-interacting factors yielded a conventionally significant effect of Condition ($t = 2.07$) in an analysis with random condition slopes for subjects and items.

A similar result was obtained for the proportion of regressions out. A logistic regression model (Jaeger, 2008) was used to analyze these proportions with random subject and item slopes. The effect of Condition did not approach significance in either Region 2 or 3 ($t < 1.0$). However, an analysis that combined Condition and Regions 4 and 5 as non-interacting factors did result in a significant effect of Condition ($z = 2.25, p < .03$): Regressions were more frequent in the NAI than the non-NAI conditions.

Discussion

Measures that reflected re-reading (proportion of regressions out, go-past time) provided clear evidence that the final clause of the first sentence disrupted reading when it failed to address the apparent QUD (in the NAI condition). This disruption was not immediate, showing up in part on the sentence that followed this final clause. There was no clear evidence that readers treated the denial (Region 2) as an ellipsis (as they may have done in Experiment 2) or that their reading was disrupted by a clash between the denial and the initial clause of the non-NAI condition. Thus, Experiment 3 provides evidence for the interpretation we offered in connection with Experiment 2: the final clause of the sentence does not address the QUD introduced by the initial clause in the NAI condition, and therefore its reading is disrupted. We do not claim that the alternate interpretation of Experiment 2 has no validity. Some of the difference in reading time seen in that experiment may have been due to the effects of interpreting the denial phrase as elliptical. But Experiment 3 in combination with Experiment 2 provides evidence for our initial claim, that discourses are easier to integrate when they address previously-established QUD's.

General Discussion

Three experiments measured subjects' eye movements while they read short discourses. The first demonstrated that reading was disrupted when an answer to an explicit question failed to address the presumed content of the question, as biased by the subcategorization preference of the verb of the question. The second demonstrated that reading was disrupted when material towards the end of the discourse failed to address the QUD introduced by the initial clause. The third provided evidence supporting this interpretation and ruling out an alternative interpretation that claimed identifying the denial phrase as elliptical was the whole story behind the Experiment 2 effects. Both phenomena were interpreted in terms of a 'Question under Discussion' (QUD) (Roberts, 1996/to appear, 2004). In the first experiment, the QUD was introduced by an explicit question (more precisely, by the most likely interpretation of an explicit question). In the second and third experiments, the QUD was introduced by an implicature (specifically, a non-actuality implicature, or NAI). Both

experiments showed that the course of normal reading was affected by the congruence between the presumed QUD and the content of the sentence.

Sources of the QUD

Overt questions—We assume that when an overt question is present, it IS the QUD. In simple discourses of the type we have investigated, this assumption seems unproblematic. However, the assumption is not invariably right, e.g., with rhetorical questions (where only one alternative is already assumed to hold), or in certain rich contexts. The semantic denotation of the question is the set of its possible or true answers, a set of alternatives (Hamblin, 1973). The following utterance in the ideal case will directly answer the question by eliminating all alternatives except one. When a complete answer is not possible, an incomplete but still coherent answer may consist of simply narrowing the set of alternatives or assigning them probabilities, or simply offering information that might help the questioner infer an answer to the question.

Focus—Focus semantics characterizes contrastive focus as introducing or presupposing alternatives. A contrastively focused constituent eliminates alternatives; the full set of alternatives is termed the *focus value* of the constituent. E.g., if the domain of discourse is Tom, Dick and Harry, then the focus value of the accented *Tom* in (4a) is the set of alternatives in (4b) (see Rooth 1985).

(4) a. TOM kissed Marcie.

b. {Dick kissed Marcie, Harry kissed Marcie, Tom kissed Marcie.}

The contrastively focused constituent is presumably most easily interpreted if preceding context introduced the set of alternatives – i.e., a QUD that the contrastively focused constituent addresses.

Focus can also guide the interpretation of ellipsis. Consider (5). Given the focus on *Sue*, Speaker B's reply is most likely interpreted as Greg also thought it was Sue who went to the market. This is because the focus on *Sue* potentially introduces alternatives to who Bill thought went to the market, as well as to who did go to the market). These two sets of alternatives constitute distinct QUDs. We have assumed that a QUD is preferably based on the main assertion of a sentence, which in this case involves the identity of who was *thought* to go to the market. Thus, Speaker B's response should be taken to address this QUD by being about who Greg thought went to the market. Note that the bias toward interpreting *Greg did too* as meaning that Greg thought Sue went to the market seems to be substantially less if *Sue* is not focused, and thus does not introduce a set of focus alternatives.

(5) Speaker A: John thought Mary went to the market. But, Bill thought SUE went to the market.

Speaker B: Greg did too.

Disjunction—Alternatives besides the formal alternatives introduced by questions and by focus can be introduced by the linguistic representation of the sentence. We have presented evidence about one such way of introducing alternatives, namely, the alternatives introduced by an NAI. As another example, consider disjunctions. Disjunctions introduce alternatives. The alternatives in (6a) constitute a QUD, which following material should address. The sentence *Sam did too* in (6a) does not address it, and seems slightly incongruous. This is not the case in (6b), which lacks the QUD of (6a).

- (6). a. John or Bill went to town. Sam did too.
 b. John and Bill went to town. Sam did too.

Other sources of QUDs—We have emphasized how discourse coherence can be increased when a prior utterance introduces one or a small set of likely QUDs that can be addressed by a following utterance. In the simplest case, the alternatives arise as an automatic consequence of computing the syntactic structure and compositional interpretation of the preceding utterance, e.g., alternatives in the semantic denotation of questions or alternatives triggered during the computation of the focus value of contrastive focus, and alternatives created as a by-product of computing independently-warranted conversational implicatures. When the interpretation of the prior utterance does not involve computing alternatives, the structure of the utterance, together with the likely position of informational focus (Carlson et al., 2009) still allows a somewhat likely QUD to be identified, by identifying the likely focus of the main assertion, and for its content substituting a variable that ranges over the contextually possible alternatives.

However, we acknowledge that sometimes discourse addresses only extremely general questions, such as ‘What is the way things are’ (Roberts, 1996/To appear), and we acknowledge that sometimes the relevance of a given utterance to preceding discourse is only apparent when the QUD the utterance addresses is identified after the fact.

Finally, we acknowledge that a reader or listener may have to be adequately motivated to add relevant QUDs to the interpretation of texts. Sometimes a QUD, a set of alternatives, is introduced by the ordinary compositional semantic interpretation of a phrase or sentence; sometimes it is introduced inferentially, as we have discussed in connection with NAIs; and sometimes it is introduced pragmatically (e.g., an indefinite noun phrase may pragmatically presuppose alternatives; the phrase *a paper* in the sentence *The editor accepted a paper* suggests that there may have been other papers that may not have been accepted). Readers and listeners may require different levels of involvement to introduce QUDs in these different situations.

Conclusions—We have provided evidence that reading of discourse is facilitated when sentences in the discourse directly address Questions under Discussion (QUDs) introduced by prior discourse. Experiment 1 showed this for explicit questions that introduced QUDs., Experiments 2 and 3 showed that QUDs could be introduced implicitly, by a type of implicature we called a Non-Actuality Implicature (NAI)

In more general terms, we have proposed some very simple but general principles for how the relation between QUDs and their answers may help to constrain the interpretation of discourse, and thereby facilitate discourse integration.. We have sketched some possible extensions of our analysis of QUDs by discussing some specific instances where particular QUDs are likely and their likely effects on the interpretation of following discourse. While we recognize that our proposals are preliminary and incomplete, we submit that they have wide applicability in the analysis of discourse coherence, and that it is well worth investigating their relation to other principles of discourse coherence discussed earlier in this paper.

Acknowledgments

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. We are particularly grateful to Nadine Bader, Sigrid Beck, Seth Cable, Brady Clark, Jesse

Harris, Chris Kennedy, and Malte Zimmerman for discussion of the issues addressed, and to Jerome Myers and Adrian Staub for their critical comments on earlier versions of this paper.

Appendix: Materials used in Experiment 1

ITEMS 1- 20: ANSWERS WITH NP BIAS VERBS

1a,b What did the talented photographer maintain?

He accepted the proposal with much reluctance.

He accepted the proposal could not be adopted.

1c,d What did the talented photographer decide?

He accepted the proposal with much reluctance.

He accepted the proposal could not be adopted.

2a, b What did the newspaper editor emphasize?

He advocated the truth publicly and with great zeal.

He advocated the truth needed to be made public.

2c,d What did the newspaper editor suggest?

He advocated the truth publicly and with great zeal.

He advocated the truth needed to be made public.

3a, b What did the CIA director write?

He confirmed the rumor once it had spread widely.

He confirmed the rumor should have been stopped sooner.

3c, d What did the CIA director say?

He confirmed the rumor once it had spread widely.

He confirmed the rumor should have been stopped sooner.

4a,b What did the angry father reveal?

He emphasized the problems faced by those who lived far away.

He emphasized the problems were continuing to get worse.

4c,d What did the angry father declare?

He emphasized the problems faced by those who lived far away.

He emphasized the problems were continuing to get worse.

5a,b What did the primary investigator reveal?

She established the alibi with the help of an old spy.

She established the alibi had been a total lie.

5c,d What did the primary investigator prove?

She established the alibi with the help of an old spy.

She established the alibi had been a total lie.

6a,b What did the gossip neighbor confirm?

She heard the story of how the five others had left.

She heard the story had never actually been true.

6c,d What did the gossip neighbor conclude?

She heard the story of how the five others had left.

She heard the story had never actually been true.

7a,b What did the confident engineer accept?

He maintained the ideas of his predecessor.

He maintained the ideas of his predecessor were outdated.

7c,d What did the confident engineer believe?

The confident engineer maintained the ideas of his predecessor.

The confident engineer maintained the ideas of his predecessor were outdated.

8a,b What did the journal editor write?

She printed the article with the footnotes at the end.

She printed the article had been too controversial.

8c,d What did the journal editor report?

She printed the article with the footnotes at the end.

She printed the article had been too controversial.

9a,b What did the frustrated tourists confirm?

They understood the message after it had been explained.

They understood the message would mean they couldn't go.

9c,d What did the frustrated tourists realize?

They understood the message after it had been explained.

They understood the message would mean they couldn't go.

10a,b What did the art critic reveal?

She wrote the interview with little regard for readers.

She wrote the interview had been a complete disaster.

10c,d What did the art critic report?

She wrote the interview with little regard for readers.

She wrote the interview had been a complete disaster.

11a,b What did the insurance investigator reveal?

She confirmed the report of the claimant.

She confirmed the claimant was crippled.

11c,d What did the insurance investigator believe?

She confirmed the report of the claimant.

She confirmed the claimant was crippled.

12a,b What did the journalist emphasize?

She grasped the dire situation of the immigrants.

She grasped the situation of the immigrants was dire.

12c,d What did the journalist claim?

She grasped the dire situation of the immigrants.

She grasped the situation of the immigrants was dire.

13a,b What did the curious inspector write?

He confirmed the blackmail rumor concerning the mayor.

He confirmed the mayor was being blackmailed.

13c,d What did the curious inspector say?

He confirmed the blackmail rumor concerning the mayor.

He confirmed the mayor was being blackmailed.

14a,b What did the supervisor reveal?

He emphasized the problems created by delays.

He emphasized the problems were due to delays.

14c,d What did the supervisor declare?

He emphasized the problems created by delays.

He emphasized the problems were due to delays.

15a,b What did the defendant's lawyer reveal?

He established the defendant's true identity.

He established the defendant had a hidden identity.

15c,d What did the defendant's lawyer prove?

He established the defendant's true identity.

He established the defendant had a hidden identity.

16a,b What did the nosy shopkeeper grasp?

She heard the story about the visiting relatives.

She heard the story was made up by the visiting relatives.

16c,d What did the nosy shopkeeper believe?

She heard the story about the visiting relatives.

She heard the story was made up by the visiting relatives.

17a,b What did the talented administrator advise?

She maintained the public's proposal.

She maintained the public proposal was correct.

17c,d What did the talented administrator think?

She maintained the public's proposal.

She maintained the public proposal was correct.

18a,b What did the cub scout reveal?

He wrote the dangerous diary during the camping trip.

He wrote the camping trip was dangerous.

18c,d What did the cub scout report?

He wrote the dangerous diary during the camping trip.

He wrote the camping trip was dangerous.

19a,b What did the club president confirm?

He understood the message about the new regulations.

He understood the message signified new regulations.

19c,d What did the club president know?

He understood the message about the new regulations.

He understood the message signified new regulations.

20a,b What did the French baker maintain?

She wrote the recipe with little regard for readers.

She wrote the recipe had been a complete triumph.

20c,d What did the French baker declare?

She wrote the recipe with little regard for readers.

She wrote the recipe had been a complete triumph.

ITEMS 21-40: ANSWERS WITH S-BIASED VERBS

21a,b What did the divorce lawyer advocate?

She argued the issue with her colleague across the state.

She argued the issue was irrelevant to the case.

21c,d What did the divorce lawyer conclude?

She argued the issue with her colleague across the state.

She argued the issue was irrelevant to the case.

22a,b What did the job applicant establish?

She believed the interviewer when she discussed things with her.

She believed the interviewer had been dishonest with her.

22c,d What did the job applicant realize?

She believed the interviewer when she discussed things with her.

She believed the interviewer had been dishonest with her.

23a,b What did the weary traveler maintain?

He claimed the proposal as his own.

He claimed the proposal was his own.

23c,d What did the weary traveler suggest?

He claimed the proposal as his own.

He claimed the proposal was his own.

24a,b What did the overworked principal write?

She concluded the proposal with a joke about style.

She concluded the proposal had not gone very well.

24c,d What did the overworked principal declare?

She concluded the proposal with a joke about style.

She concluded the proposal had not gone very well.

25a,b What did the bank guard establish?

He confessed the robbery after a sleepless night.

He confessed the robbery had been his own idea.

25c,d What did the bank guard declare?

He confessed the robbery after a sleepless night.

He confessed the robbery had been his own idea.

26a,b What did the experienced judge maintain?

He decided the appeal on the merit of the case.

He decided the appeal should be started right away.

26c,d What did the experienced judge suggest?

He decided the appeal on the merit of the case.

He decided the appeal should be started right away.

27a,b What did the shrewd salesman establish?

He figured the prices for the recent book sale.

He figured the prices would be going up soon.

27c,d What did the shrewd salesman suggest?

He figured the prices for the recent book sale.

He figured the prices would be going up soon.

28a,b What did the novice plumber understand?

He realized the mistake before calling the owner.

He realized the mistake would cost someone some money.

28c,d What did the novice plumber confess?

He realized the mistake before calling the owner.

He realized the mistake would cost someone some money.

29a,b What did the film director advise?

She suggested the scene that the writer removed.

She suggested the scene should be filmed at night.

29c,d What did the film director believe?

She suggested the scene that the writer removed.

She suggested the scene should be filmed at night.

30a,b What did the high school principal maintain?

He suspected the teenager before hearing the facts.

He suspected the teenager wasn't telling the truth.

30c,d What did the high school principal believe?

He suspected the teenager before hearing the facts.

He suspected the teenager wasn't telling the truth.

31a,b What did the accomplished actress maintain?

She proved the difficulty of satisfying the director.

She proved the difficulty could eventually be overcome.

31c,d What did the accomplished actress decide?

She proved the difficulty of satisfying the director.

She proved the difficulty could eventually be overcome.

32a,b What did the new boss understand?

He guessed the employee was secretive about his private life.

He guessed the employee secrets about his private life.

32c,d What did the new boss realize?

He guessed the employee was secretive about his private life.

He guessed the employee secrets about his private life.

33a,b What did the tired customs official establish?

He doubted the declaration on the traveler's landing form

He doubted the declaration was completely accurate.

33c,d What did the tired customs official suggest?

He doubted the declaration on the traveler's landing form

He doubted the declaration was completely accurate.

34a,b What did the hardworking principal decide?

He concluded the proposal with a disclaimer.

He concluded the proposal was weak.

34c,d What did the overworked principal realize?

He concluded the proposal with a disclaimer.

He concluded the proposal was weak.

35a,b What did the parking attendant establish?

He confessed the damage to the Mercedes.

He confessed the damage to the Mercedes was his fault.

35c,d What did the parking attendant declare?

He confessed the damage to the Mercedes.

He confessed the damage to the Mercedes was his fault.

36a,b What did the trial judge maintain?

He decided the case strictly on its merits.

He decided the case simply had no merit.

36a,b What did the trial judge suggest?

He decided the case strictly on its merits.

He decided the case simply had no merit.

37a,b What did the sneaky mechanic confirm?

He figured the high estimate for the brakes.

He figured the break estimate would be high.

37c,d What did the sneaky mechanic suggest?

He figured the high estimate for the brakes.

He figured the brake estimate would be high.

38a,b What did the experienced carpenter understand?

He realized the oversight in the design.

He realized the design oversight could be dangerous.

38c,d What did the experienced carpenter suspect?

He realized the oversight in the design.

He realized the design oversight could be dangerous.

39a,b What did the loud caterer emphasize?

She suggested the grilled spiced meat

She suggested the meat should be grilled.

39c,d What did the loud caterer declare?

She suggested the grilled spiced meat

She suggested the meat should be grilled.

40a,b What did the fish vendor maintain?

She suspected the old eel.

She suspected the eel was old.

40c,d What did the fish vendor believe?

She suspected the old eel.

She suspected the eel was old.

Materials Used in Experiment 2

1 Mary could visit me, but she hasn't, according to what I hear, visited her parents as of yet. Mary visited me, but she hasn't, according to what I hear, visited her parents as of yet.

2 Jessica could go to Paris, but she hasn't, according to her sister, been to Rome any time recently.

Jessica went to Paris, but she hasn't, according to her sister, been to Rome any time recently.

3 Caitlin could ace all her exams, but she hasn't, as far as I know, opened a book all semester. Caitlin aced all her exams, but she hasn't, as far as I know, opened a book all semester.

4 Max could write poetry, but he hasn't, to our knowledge, even kept a journal or written much else.

Max wrote poetry, but he hasn't, to our knowledge, even kept a journal or written much else.

5 Josh could plan a round-the-world trip, but he hasn't, according to his buddies, asked for vacation time at his job.

Josh planned a round-the-world trip, but he hasn't, according to his buddies, asked for vacation time at his job.

6 Dan could date this Vogue model, but he hasn't, for some reason, taken her to dinner at a nice restaurant.

Dan dated this Vogue model, but he hasn't, for some reason, taken her to dinner at a nice restaurant.

7 Theo could travel around the country, but he hasn't, according to his wife, ever found time for a real vacation.

Theo traveled around the country, but he hasn't, according to his wife, ever found time for a real vacation.

8 Lisa could play Mafia Wars, but she won't, according to her roommate, friend enough people to win.

Lisa played Mafia Wars, but she won't, according to her roommate, friend enough people to win.

9 Simon could excel at work, but he won't, we suspect, ever put in the necessary hours to be promoted.

Simon excelled at work, but he won't, we suspect, ever put in the necessary hours to be promoted.

10 Jordana could maintain a pretty good grade average, but she won't, no matter how much we push, give up drinking with her friends.

Jordana maintained a pretty good grade average, but she won't, no matter how much we push, give up drinking with her friends.

11 Tina could get a job in New York, but she hasn't, as far as we know, persuaded her husband to move.

Tina got a job in New York, but she hasn't, as far as we know, persuaded her husband to move.

12 Ruth could take care of the kids for us often, but she hasn't, as far as I know, given us an answer about next week.

Ruth took care of the kids for us often, but she hasn't, as far as I know, given us an answer about next week.

13 Martin could give a guest lecture, but he won't, I'm told, agree to grade any term papers. Martin gave a guest lecture, but he won't, I'm told, agree to grade any term papers.

14 Trisha could write to the owner, but she didn't, as far as we know, request a raise for herself. Trisha wrote to the owner, but she didn't, as far as we know, request a raise for herself.

15 Max could become an astronaut, but he won't, according to his mother, give up mathematics, his true love.

Max became an astronaut, but he won't, according to his mother, give up mathematics, his true love.

16 Kevin could fix his car, but he won't, according to his brother, buy the necessary parts because they're too expensive.

Kevin fixed his car, but he won't, according to his brother, buy the necessary parts new because they're too expensive.

17 Wendell could pick Mom up at the airport, but he hasn't, we all know, cleaned out the car for the trip.

Wendell picked Mom up at the airport, but he hasn't, we all know, cleaned out the car for the trip.

18 Irwin could persuade Lucy, but he didn't, according to Jim, try to persuade Sue about anything.

Irwin persuaded Lucy, but he didn't, according to Jim, try to persuade Sue about anything.

19 John could teach programming, but he hasn't, we suspect, prepared any assignments for the students.

John could teach programming, but he hasn't, we suspect, prepared any assignments for the students.

20 Lulu could watch Sesame Street, but she hasn't, according to the babysitter, watched Sponge Bob for the past few weeks.

Lulu watched Sesame Street, but she hasn't, according to the babysitter, watched Sponge Bob for the past few weeks.

Materials used in Experiment 3

Non-NAI then NAI forms 1. This information was released but Gorbachev didn't, according to our sources, notify the appropriate bureau. That may cause problems later on.

This information should be released but Gorbachev didn't, according to our sources, notify the appropriate bureau. That may cause problems later on.

2. A trip home was planned for August but Anne didn't, as far as we know, invite her parents for Christmas. She'll probably get into trouble for that.

A trip home should be planned for August but Anne didn't, as far as we know, invite her parents for Christmas. She'll probably get into trouble for that.

3A joint meeting was scheduled but the director didn't, to our knowledge, inform his agency. He doesn't want the agency to tell him what to do.

A joint meeting should be scheduled but the director didn't, to our knowledge, inform his agency. He doesn't want the agency to tell him what to do.

4. A private firm was hired but the director didn't, according to the record books, approve the expenditure. It seems that he doesn't really want the work to be done.

A private firm should be hired but the director didn't, according to the books, approve the expenditure. It seems that he doesn't really want the work to be done.

5. An urban planner was consulted but the supervisor didn't, as far as anybody could tell, even put the topic on the agenda. The supervisor is apparently trying to cancel the project.

An urban planner should be consulted but the supervisor didn't, as far as anybody could tell, even put the topic on the agenda. The supervisor is apparently trying to cancel the project.

6. A hybrid car was ordered but the energy expert didn't, according to the latest report, project the gas mileage. It's not clear how big the actual savings will be.

A hybrid car was supposed to be ordered but the energy expert didn't, according to the latest report, project the gas mileage. It's not clear how big the actual savings will be.

7. An expansion was planned for the branch office but the home office didn't, according to office gossip, approve the expense. They want to keep costs way down.

An expansion was supposed to be planned for the branch office but the home office didn't, according to office gossip, approve the expense. They want to keep costs way down.

8. A taxi driver was called but Sally didn't, as far as we can tell, get ready to leave. She's certainly going to be late.

A taxi driver was supposed to be called but Sally didn't, as far as we can tell, get ready to leave. She's certainly going to be late.

9. A planning committee was put together but the CEO didn't, as of yesterday, hire any additional secretarial help. He may be secretly opposed to the plans.

A planning committee was supposed to be put together but the CEO didn't, as of yesterday, hire any additional secretarial help. He may be secretly opposed to the plans.

10. This research proposal was revised but the head investigator didn't, according to the grants office, redo the budget. The investigator doesn't have much sense about money.

This research proposal was supposed to be revised but the head investigator didn't, according to the grants office, redo the budget. The investigator doesn't have much sense about money.

11. The cookies were made but the babysitter didn't, according to the kids, even give them milk. The kids were going to be pretty unhappy.

The cookies should be made but the babysitter didn't, according to the kids, even give them milk. The kids were going to be pretty unhappy.

12. The house was sold but the new agent didn't, according to my source, advertise in the usual papers. He didn't even hang up a sign.

The house should be sold but the new agent didn't, according to my source, advertise in the usual papers. He didn't even hang up a sign.

13. A painting was donated but the usual donor didn't, even after much prompting, contribute money for the fund raising drive. He may be feeling short of cash.

A painting should be donated but the usual donor didn't, even after much prompting, contribute money for the fund raising drive. He may be feeling short of cash.

14. A financial advisor was contacted but the director didn't, according to the secretary, say anything about his suspicions. He has reasons to be pretty suspicious.

A financial advisor should be contacted but the director didn't, according to the secretary, say anything about his suspicions. He has reasons to be pretty suspicious.

15. This oriental rug was repaired but the original seller didn't, according to the current owner, send it out to qualified experts. He probably didn't realize how valuable it was.

This oriental rug should be repaired but the original seller didn't, according to the current owner, send it out to qualified experts. He probably didn't realize how valuable it was.

16. The new employee was trained but the manager didn't, as far as we see, leave enough time to do it right. That's going to give the new guy some big problems later on.

The new employee is supposed to be trained but the manager didn't, as far as we see, leave enough time. That's going to give the new guy some big problems later on.

17. A deck was added to the cottage but the tenant didn't, according to the realtor, want the construction to take place when he was home. He didn't want to hear all the noise.

A deck is supposed to be added to the cottage but the tenant didn't, according to the realtor, want the construction to take place when he was home. He didn't want to hear all the noise.

18. This tractor was inspected but the mechanic didn't, according to the supervisor, know what he expected to find. The mechanic wasn't very experienced in tractor repair.

This tractor is supposed to be inspected but the mechanic didn't, according to the supervisor, know what he expected to find. The mechanic wasn't very experienced in tractor repair.

19. A new strategy was devised but the director didn't, we know for a fact, contribute any ideas at all. He really doesn't seem to do much of anything worthwhile.

A new strategy is supposed to be devised but the director didn't, we know for a fact, contribute any ideas at all. He really doesn't seem to do much of anything worthwhile.

20. A diplomatic exchange was arranged but the ambassador didn't, according to the newspaper, contact the cultural attache. This may be a breach of protocol.

A diplomatic exchange is supposed to be arranged but the ambassador didn't, according to the newspaper, contact the cultural attache. This may be a breach of protocol.

References

- Arregui A, Clifton C Jr, Frazier L, Moulton K. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language*. 2006; 55:232–246. [PubMed: 17710192]
- Baayen, RH. *Analyzing linguistic data*. Cambridge University Press; Cambridge: 2008.
- Bates DM. Fitting linear mixed models in R: Using the lme4 package. *R News: The Newsletter of the R Project*. 2005; 5(1):27–30.
- Beaver, DL.; Clark, B. *Sense and sensitivity: How focus determines meaning*. Blackwell; Oxford: 2008.
- Branigan HP, Pickering MJ, Liversedge SP, Stewart AJ, Urbach TP. Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research*. 1995; 24:489–506.
- Branigan HP, Pickering M, McLean JF. Priming prepositional-phrase attachment during comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 2005; 31:468–481.
- Clifton C Jr, Frazier L, Connine C. Lexical expectations in sentence comprehension. *Journal of Verbal Learning & Verbal Behavior*. 1984; 23(6):696–708.

- Crain, S.; Steedman, M. On not being led up the garden path: The use of context by the psychological parser. In: Dowty, D.; Karttunen, L.; Zwicky, A., editors. *Natural language parsing*. Cambridge University Press; Cambridge: 1985. p. 320-358.
- Crawley, MJ. *The R Book*. John Wiley & Sons; Chichester, England: 2007.
- Cuetos F, Mitchell DC. Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*. 1988; 30:73-105. [PubMed: 3180704]
- Ford, M.; Bresnan, J.; Kaplan, R. A competence-based theory of syntactic closure. In: Bresnan, J., editor. *The Mental Representation of Grammatical Relations*. MIT Press; Cambridge, MA: 1983.
- Ferguson H, Sanford A. Anomalies in real and counterfactual worlds: An eye movement investigation. *Journal of Memory & Language*. 2008; 58:609-626.
- Ferreira F, Henderson J. The use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1990; 16:555-568.
- Frazier, L. Sentence processing: A tutorial review. In: Coltheart, M., editor. *Attention and performance*. Lawrence Erlbaum Associates; Hillsdale, NJ: 1987. p. 559-586.
- Frazier, L. Against lexical generation of syntax. In: Marslen-Wilson, William, editor. *Lexical Representation and Process*. Bradford Books (MIT Press); Cambridge: 1989.
- Frazier L, Clifton C Jr. The syntax-discourse divide: Processing ellipsis. *Syntax*. 2005; 8:154-207.
- Frazier L, Taft L, Roeper T, Clifton C Jr. Ehrlich K. Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition*. 1984; 12:421-430.
- Gahl S, Garnsey S. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*. 2004; 80(4):748-775.
- Garnsey SM, Pearlmutter NJ, Myers E, Lotocky MA. The contributions of verb bias and plausibility to the comprehension of temporarily-ambiguous sentences. *Journal of Memory and Language*. 1997; 37:58-93.
- Garrod, SC.; Sanford, AJ. Resolving sentences in a discourse context: How discourse representation affects language understanding. In: Gernsbacher, MA., editor. *Handbook of Psycholinguistics*. Academic Press; San Diego: 1994.
- Gernsbacher MA. Two decades of structure building. *Discourse Processes*. 1997; 23:265-304.
- 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]
- Hamblin C. Questions in Montague English. *Foundations of Language*. 1973; 10:41-53.
- Jaeger TF. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*. 2008; 59:434-446. [PubMed: 19884961]
- Johnson, K. What VP ellipsis can do, and what it can't, but not why. In: Baltin, M.; Collins, C., editors. *The handbook of contemporary syntactic theory*. Blackwell; Oxford: 2001. p. 439-479.
- Kamide Y. Anticipatory processes in sentence processing. *Language and Linguistics Compass*. 2008; 2:647-640.
- Karttunen L. Syntax and semantics of questions. *Linguistics and Philosophy*. 1977; 1:3-44.
- Kennison SM. Limitations on the use of verb information during sentence comprehension. *Psychonomic Bulletin & Review*. 2001; 8:132-137. [PubMed: 11340858]
- McCrudden, MT.; Magliano, JP.; Schraw, G., editors. *Text relevance and learning from text*. Information Age Publishing; Charlotte, NC: 2011.
- O'Brien, EJ.; Myers, JL. Text comprehension: A view from the bottom up. In: Goldman, SR.; Graesser, AC., editors. *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso*. Lawrence Erlbaum Associates; Mahwah, NJ: 1999.
- R_Development_Core_Team. *A language and environment for statistical computing*. R Foundation for Statistical Computing; Vienna, Austria: 2007.
- Rayner K, Carlson M, Frazier L. The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*. 1983; 22:358-374.
- Rayner K, Sereno S, Morris R, Schmauder R, Clifton CJ. Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*. 1989; 4:SI 21-50.

- Roberts, C. Information structure in discourse: towards an integrated theory of pragmatics. Papers in Semantics: OSU Working Papers in Linguistics; Columbus, OH: Department of Linguistics, The Ohio State University; 1996. p. 49 to appear (To appear, Semantics and Pragmatics)
- Roberts, C. Discourse context in dynamic interpretation. In: Horn, L.; Ward, G., editors. Handbook of Contemporary Pragmatic Theory. Blackwell; Oxford: 2004. p. 197-220.
- Rohde, H.; Kehler, A. QUD-driven expectations in discourse interpretation. Paper presented at the 83rd Annual Meeting of the Linguistic Society of America; Pittsburgh, PA. 2009.
- Rohde H, Levy R, Kehler A. Anticipating explanations in relative clause processing. *Cognition*. 2011; 118:339–358.
- Sanford, A.J.; Garrod, S. Referential processes in reading: Focusing on roles and individuals. In: Balota, D.A.; d'Arcais, G.B.F.; Rayner, K., editors. Comprehension processes in reading. Erlbaum; Hillsdale, NJ: 1990. p. 465-485.
- Trabasso T, van den Broek P. Causal thinking and the representation of narrative events. *Journal of Memory and Language*. 1985; 24:612–628.
- Trueswell JC, Tanenhaus MK, Kello C. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1993; 19:528–553.
- Zondervan, A.; Meroni, L.; Gualmini, A. Experiments on the role of the question under discussion for ambiguity resolution and implicature computation in adults. Paper presented at the SALT XVIII; Ithaca, NY. 2008.
- Zwaan RA. Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 1996; 22:1196–1207.
- Zwaan, R.A.; Rapp, D.N. Discourse comprehension. In: Traxler, M.J.; Gernsbacher, M.A., editors. Handbook of psycholinguistics. 2nd ed. Elsevier; Amsterdam: 2006.

Figure 1.

Go-Past Time, ms, Experiment 1; left two columns, NP-biased verb in question; right two columns; S-biased verb in question. M = answer form matches, MM = answer form mismatches, bias of verb in question (match = NP complement for NP biased question verb, S complement for S biased question verb)

Figure 2.

Regression out proportions, Experiment 1; left two columns, NP-biased verb in question; right two columns; S-biased verb in question. M = answer form matches, MM = answer form mismatches, bias of verb in question (match = NP complement for NP biased question verb, S complement for S biased question verb)

Table 1

Illustration of items used in Experiment 1. Condition indicates syntactic bias of verb used in question and in answer, and syntactic resolution as direct object or sentence complement.

Condition			Question	Answer
Q	A	Syntax		
Bias	Bias			
NP	NP	NP	What did the primary investigator reveal?	She established the alibi with the help of an old spy.
NP	NP	S	"	She established the alibi had been a total lie.
S	NP	NP	What did the primary investigator prove?	She established the alibi with the help of an old spy.
S	NP	S	"	She established the alibi had been a total lie.
NP	S	NP	What did the job applicant establish?	She believed the interviewer when she discussed things with her.
NP	S	S	"	She believed the interviewer had been dishonest with her.
S	S	NP	What did the job applicant realize?	She believed the interviewer when she discussed things with her.
S	S	S	"	She believed the interviewer had been dishonest with her.

Table 2

Mean eyetracking measures (first pass time, ms; go-past time, ms; percentage of regressions out), Experiment 1, by condition (question bias X answer bias X syntactic resolution) and by sentence region. (Standard errors in parentheses)

Quest	Answer Syntax		Region			
	Bias	Bias	1 (subj+verb)	2 (DP)	3 (disambig)	4 (following)
			First Pass Time (ms)			
NP	NP	NP	426 (16)	462 (23)	388 (16)	556 (29)
NP	NP	S	411 (15)	448 (21)	448 (20)	497 (27)
S	NP	NP	420 (15)	463 (22)	415 (19)	543 (25)
S	NP	S	425 (16)	471 (22)	423 (17)	508 (29)
NP	S	NP	401 (14)	446 (21)	540 (20)	428 (25)
NP	S	S	410 (14)	411 (20)	426 (17)	448 (27)
S	S	NP	419 (17)	413 (18)	508 (22)	396 (21)
S	S	S	386 (13)	430 (19)	458 (18)	440 (23)
			Go-Past Time (ms)			
NP	NP	NP	504 (32)	588 (32)	513 (28)	1252 (81)
NP	NP	S	438 (19)	560 (29)	772 (63)	1128 (72)
S	NP	NP	469 (24)	631 (47)	586 (48)	1102 (69)
S	NP	S	489 (30)	555 (28)	630 (41)	1113 (83)
NP	S	NP	460 (25)	565 (26)	825 (45)	1036 (79)
NP	S	S	459 (22)	541 (35)	650 (38)	812 (64)
S	S	NP	512 (30)	542 (25)	884 (56)	1157 (112)
S	S	S	429 (20)	551 (29)	609 (40)	811 (56)
			Regressions Out (percentage)			
NP	NP	NP	3 (.015)	13 (.025)	16 (.027)	49 (.039)
NP	NP	S	2 (.011)	13 (.024)	24 (.031)	51 (.043)
S	NP	NP	3 (.013)	13 (.024)	17 (.027)	47 (.038)
S	NP	S	4 (.014)	10 (.022)	18 (.028)	52 (.043)
NP	S	NP	4 (.014)	18 (.028)	24 (.030)	50 (.043)
NP	S	S	4 (.015)	14 (.026)	22 (.030)	38 (.041)
S	S	NP	5 (.017)	19 (.028)	30 (.033)	48 (.042)
S	S	S	3 (.013)	13 (.024)	15 (.025)	39 (.041)

Table 3

Illustration of material used in Experiment 2, with analysis regions indicated by /

Condition	Example
NAI	Mary could visit me, / ₁ but she hasn't, / ₂ according to what I hear, / ₃ visited her parents / ₄ as of yet./ ₅
no-NAI	Mary visited me, / ₁ but she hasn't, / ₂ according to what I hear, / ₃ visited her parents / ₄ as of yet. / ₅

Table 4

Mean eyetracking measures, Experiment 2, by condition and sentence region. (Standard errors in parentheses)

Condition	Region			
	2 (denial)	3 (paren)	4 (critical)	5 (spillover)
First Pass Time (ms)				
NAI	423 (10)	488 (12)	592 (15)	475 (15)
non-NAI	422 (10)	477 (12)	571 (15)	456 (14)
Total Time (ms)				
NAI	476 (12)	588 (17)	781 (14)	574 (17)
non-NAI	510 (13)	581 (15)	712 (20)	531 (15)
Second Pass Time (ms) (subject means)				
NAI	56 (8)	73 (11)	161 (17)	---
non-NAI	83 (10)	42 (7)	128 (13)	---
Regressions Out (%)				
NAI	1 (1)	7 (1)	5 (1)	41 (3)
non-NAI	3 (1)	15 (2)	5 (1)	37 (2)
Go-past Time (ms)				
NAI	452 (15)	551 (15)	665 (20)	900 (43)
non-NAI	457 (12)	605 (18)	623 (16)	789 (33)

Table 5

Example of material used in Experiment 3, with line breaks and analysis regions indicated

Condition	Example
NAI	This information should be released/ ¹ but Gorbachev didn't/ ² according to our sources/ ³ notify the appropriate bureau. ⁴ That may cause problems later on. ⁵
non-NAI	This information was released/ ₁ but Gorbachev didn't/ ₂ according to our sources/ ₃ notify the appropriate bureau. ₄ That may cause problems later on. ₅

Table 6

Mean eyetracking measures, Experiment 3, by condition and sentence region. (Standard errors in parentheses)

Condition	Region				
	1 (initial)	2 (denial)	3 (paren)	4 (critical)	5 (spillover)
	First Pass Time (ms)				
NAI	1060 (30)	752 (22)	592 (21)	804 (30)	1127 (35)
non-NAI	808 (24)	766 (25)	598 (21)	797 (31)	1096 (32)
	Total Time (ms)				
NAI	1260 (41)	964 (35)	746 (25)	1003 (33)	1241 (37)
non-NAI	977 (34)	958 (31)	747 (24)	956 (33)	1157 (35)
	Second Pass Time (ms) (subject means)				
NAI	214 (30)	138 (22)	74 (11)	48 (11)	---
non-NAI	168 (22)	129 (17)	67 (11)	39 (9)	---
	Regressions Out (%)				
NAI	---	12 (2)	10 (2)	21 (2)	20 (3)
non-NAI	---	10 (2)	13 (2)	16 (2)	14 (2)
	Go-past Time (ms)				
NAI	1071 (32)	878 (28)	715 (29)	1085 (44)	1505 (70)
non-NAI	808 (24)	864 (27)	784 (39)	997 (38)	1353 (54)