Anticipating Explanations in Relative Clause Processing

H. Rohde*,a, R. Levyb, A. Kehlerb

a2016 Sheridan Road, Evanston, IL 60208-4090, USA
b9500 Gilman Drive #108, La Jolla, CA 92037-0108, USA

Abstract

In support of a model of language comprehension in which pragmatic biases are integrated with syntactic processing, we show that expectations about upcoming discourse continuations influence the resolution of local structural ambiguity. An off-line sentence-completion study and an on-line self-paced reading study examined readers’ expectations for high/low relative clause attachments following implicit-causality and non-implicit-causality verbs (John detests/babysits the children of the musician who...). In the off-line study, the widely reported low-attachment preference for English is observed in the non-implicit causality condition, but this preference gives way to more high attachments in the implicit causality condition in cases in which (i) the verb’s causally implicated referent occupies the high-attachment position and (ii) the relative clause provides an explanation for the event described by the matrix clause (e.g., ...who are arrogant and rude). In the on-line study, a similar preference for high attachment emerges in the implicit causality context – crucially, before the occurrence of any linguistic evidence that the RC does in fact provide an explanation – whereas the low-attachment preference is consistent elsewhere. These results suggest that comprehenders construct discourse contexts dynamically during the normal course of sentence interpretation, using available pragmatic cues mid-sentence to generate expectations about the structural analysis of the remainder of the sentence. This finding therefore constitutes the first demonstration that expectations about ensuing discourse relationships are as important as, and are fully integrated with, lexical and morphosyntactic cues in resolving syntactic ambiguity.

Key words: Discourse Processing; Relative Clause Attachment Ambiguity; Implicit Causality; Coherence Relations

*Corresponding author
Email address: hannah@northwestern.edu (H. Rohde)

Preprint submitted to Elsevier January 1, 2010
1. Introduction

A foundational question in psycholinguistics asks how it can be that comprehenders, working merely from linear sequences of words, are able to recover the array of latent relationships necessary to establish the speaker’s intended underlying message. This paper considers two key components of this process: the establishment of structure within the sentence (through the use of syntactic phrase-structure rules and local dependency relations) and the deduction of pragmatic relationships between sentences (through the use of world knowledge and propositional reasoning). We show that inferences about these two apparently disparate aspects of language structure are deeply interleaved.

At the sentence level, a considerable body of research has focused on the comprehender’s need to infer intrasentential structural relationships between underlying constituents in building a syntactic representation of an utterance. When processing each sentence in (1) below, for example, a comprehender infers that his coworkers stands in the structural relation of objecthood with detests, that obnoxious is coordinated with rude and unpleasant, to be around modifies unpleasant but not obnoxious or rude, and so forth.

(1) John detests his coworkers. They are obnoxious, rude, and unpleasant to be around.

Likewise, in a largely independent line of research, considerable work has also examined how relationships are recovered above the level of the sentence, that is, how comprehenders establish discourse dependencies. When processing (1), for instance, the comprehender is likely to infer an anaphoric relationship between the phrase they and the coworkers mentioned in the first sentence, and further that the second sentence stands in a particular discourse relationship with the first sentence, specifically that of providing an explanation for why John detests these coworkers.

Considerable progress has been made in our understanding of how these latent relationships are recovered at both levels. This progress, however, has produced models of sentence and discourse processing that bear strikingly little resemblance to one another, with no obvious prospect for a seamless integration. When considering processing mechanisms at both levels as an ensemble, therefore, several key questions are naturally brought to the fore:

1. Do syntactic (i.e. intrasentential) and discourse (i.e. intersentential) comprehension processes happen sequentially, or are they interleaved?
2. If they are interleaved, is there information flow from processing at the
Higher (discourse) level that influences decisions at the lower (intrasentential) level, and if so, is this influence fully incremental?

3. Is discourse information utilized during the normal course of sentence interpretation, as opposed to being invoked only when sentence grammaticality or felicity is at stake?

Under clausal-integration accounts of discourse processing (Garnham et al., 1996; Stewart et al., 2000), for instance, construction of the discourse relationship between two clauses begins only after each clause has individually been processed in full, so that the conventional answer to these questions is “no”. In this paper, we propose that the correct answer to these questions is “yes”, on the basis of two experiments focusing on the juxtaposition of a quintessentially structure-oriented, intrasentential comprehension process—the resolution of syntactic attachment of relative clauses (RCs)—with what is generally taken to be a quintessentially meaning-oriented, discourse-level comprehension process—the establishment of the discourse relationships that hold between clauses. Corresponding to the three questions just posed, we argue (1) that syntactic and discourse comprehension processes are interleaved, (2) that inferences about likely discourse relationships holding between two clauses can indeed influence syntactic processing decisions fully incrementally, i.e., before the clauses have been processed in their entirety, and (3) that these influences are present generally, and not invoked only in the face of potential sentence unacceptability. We further argue that, when cast within a suitable evidential and/or probabilistic model, both syntactic and discourse processing (as well as their interdependencies) can be naturally captured by appealing to common interpretation principles and mechanisms. To our knowledge, this is the first such demonstration of its kind.

The remainder of this section establishes the theoretical background for our approach to discourse relations, introduces the structural ambiguity of RC attachment, and presents a series of observations regarding pragmatic interpretation that together will allow us to use RC attachment ambiguity as a testbed for a model of language comprehension in which pragmatic biases regarding discourse relations are integrated with syntactic processing.

1.1. Coherence Establishment Between and Within Sentences

While there are many aspects of discourse understanding that are poorly understood, there is one thing that we can be sure of: Discourses are not simply arbitrary collections of utterances. A felicitous discourse must instead meet the criterion of being coherent. As mentioned above, comprehenders do not generally interpret the two statements in passage (1) as independent facts about John and his coworkers, but instead infer a particular coherence relation between
them – in this case, Explanation – under which the second clause is interpreted as providing the cause or reason for the state described by the first. While this relationship is by no means entailed by the passage, the assumption that such a relationship holds is easily accommodated, in light of the comprehender’s causal knowledge and assumption that some discourse relation holds between the two. If it were to subsequently become clear that John does not mind rude and unpleasant people, but instead dislikes his co-workers because of their fashion sense, example (1) no longer reads as a coherent passage.

As noted by Hobbs (1990), the idea that such coherence relations are a fundamental aspect of discourse production and understanding was recognized by Hume, who in his *Inquiry Concerning Human Understanding* (1748) states:

> Were the loosest and freest conversation to be transcribed, there would immediately be observed something which connected it in all its transitions. Or where this is wanting, the person who broke the thread of discourse might still inform you, that there had secretly revolved in his mind a succession of thought, which had gradually led him from the subject of conversation. (Section III, “Of the Association of Ideas”)

We are therefore led to ask what the inventory of these coherence relations is, or as Hume put it in his abstract of *A Treatise of Human Nature*, what the principles are that serve as “a secret tie or union among particular ideas, which causes the mind to conjoin them more frequently together, and makes the one, upon its appearance, introduce the other”. Famously, Hume found but three:

> Though it be too obvious to escape observation, that different ideas are connected together; I do not find that any philosopher has attempted to enumerate or class all the principles of association; a subject, however, that seems worthy of curiosity. To me, there appear to be only three principles of connexion among ideas, namely, Resemblance, Contiguity in time or place, and Cause or Effect. (Section III, “Of the Association of Ideas”)

As suggested in Kehler (2002), the Explanation relation we identified for (1) is naturally classified in Hume’s Cause-Effect category. Passage (2), on the other hand, does not involve causality on its most natural interpretation. Instead, it is characterized by an Elaboration relation – an instance of Hume’s Resemblance category – in which the second clause is understood to elaborate the first.
(2) John babysits Bobby every Saturday night. He hangs out with him for four hours of TV, games, and reading stories.

As a result, the comprehender infers that the two sentences jointly describe a composite event, rather than inferring that the hanging out is distinct from the babysitting.

Explanation and Elaboration are only two relations from a larger inventory that are categorized within Hume’s trichotomy by Kehler (2002). Kehler argues therein that the behavior of a variety of discourse-sensitive linguistic phenomena is dependent on which category the operative coherence relation is a member of. Recent studies by Rohde et al. (2006, 2007) and Kehler et al. (2008) provide evidence that comprehenders not only infer such coherence relations, but also generate expectations concerning what coherence relations are likely to ensue based on the current context. They further argue that any successful account of pronoun interpretation necessarily must incorporate those expectations. These recent findings contrast with the previous work of Garnham et al. (1996) and Stewart et al. (2000), which argues that coherence-biasing information affects only later stages of processing, crucially after the interpretations of pairs of clauses are fully available for integration.

1.2. Coherence Relations and Syntactic Processing

In consideration of these results, one might go one step further and ask whether discourse coherence expectations can be shown to influence not only discourse-dependent expressions such as pronouns, but also local syntactic processing decisions within a sentence. If so, it would constitute a fairly radical demonstration of the range of information sources that are brought to bear in on-line syntactic comprehension. We address this question by examining the potential influence of discourse coherence expectations in the online disambiguation of RC attachment ambiguity. RC attachment provides an ideal testing ground because, as we will demonstrate in Section 1.4, an RC is linked to its matrix clause by both a syntactic relationship and a discourse coherence relationship. A classic example of this attachment ambiguity is shown in (3):

(3) Someone shot the servant of the actress who was on the balcony.

Example (3) contains an RC (who was on the balcony) that can be interpreted to modify one or the other of the two noun phrases (NPs) in the preceding complex NP (the servant of the actress). The RC is said to attach HIGH if it is interpreted as modifying NP1 (the servant), which occupies the higher position
in the syntactic structure. It is likewise said to attach low if it modifies NP2 (the actress), the possessor NP within the complex NP. Following an early account that predicted a low-attachment preference through the principle of Late Closure (Frazier, 1978), a low-attachment preference in contexts like (3) in English is now widely accepted as the default, having been confirmed in off-line studies with questionnaires and completion tasks and in most on-line studies (Frazier and Clifton, 1996; Carreiras and Clifton, 1999; Fernandez, 2003; but see also Traxler et al., 1998). More recent research, however, has made clear that the attachment-preference picture is in fact considerably more complex and depends on lexico-syntactic details of the ambiguity in question. We defer further discussion of these issues to Section 4.1.

This previous work, however, has not explored whether (expectations about) the discourse relationship between the matrix clause and the embedded clause can influence the syntactic attachment decision. Our experiments will therefore employ cases in which the matrix and relative clauses share both a syntactic and a discourse relation, and hence in which syntactic and discourse processing might be expected to interact. Our design brings together three independent observations involving pragmatic interpretation:

i. Matrix clauses that contain so-called implicit causality (IC) verbs create a strong expectation that an ensuing clause will contain an explanation of the eventuality denoted by the matrix clause.

ii. An RC can be used to provide an explanation of the matrix clause.

iii. Certain IC verbs, specifically object-biased IC verbs, create an expectation that the locus of an ensuing explanation is likely to be the verb’s direct object.

On an evidential and/or probabilistic theory in which discourse processing occurs simultaneously with syntactic comprehension and can therefore influence on-line attachment decisions, these three factors should conspire to generate a bias toward high attachments for RCs in a particular class of ambiguous sentences. The reason involves a chain of logic that combines a variety of different insights, which we walk through in the following subsections.

1.3. Observation 1: Implicit Causality Creates an Expectation for an Explanation

Our first observation is that a clause containing a certain type of verb, of which detest in (4) below is an example, typically creates a strong expectation that it will be followed by an explanation of the eventuality that the clause denotes.¹ Garvey

¹Unless, of course, the cause of the eventuality has already been provided previously in the discourse (cf. Simner and Pickering, 2005).
and Caramazza (1974) coined the term implicit causality (IC) to describe the verbs of interest here, for which one of the role participants “is implicated as the assumed locus of the underlying cause of the action or attitude” (see Section 1.5 below). Kehler et al. (2008) conducted a story continuation experiment using IC and non-IC verbs as in (4)-(5).

(4) John detests Mary. [IC VERB]
(5) John babysits Mary. [NON-IC VERB]

Kehler et al. found that context sentences with IC verbs (4) yielded significantly more Explanation coherence continuations (60%) than context sentences with non-IC verbs (5) did (24%). At an intuitive level, the lexical semantics of verbs like detest appear to lead the comprehender to ask Why? in a way that verbs like babysit do not.

1.4. Observation 2: RCs can Provide an Explanation

Restriction of reference is only one of the possible functions of an RC. Our second observation brings another role that RCs play to the fore, one that has not, to our knowledge, previously been utilized in psycholinguistic work: An RC can serve the additional pragmatic function of providing an explanation of the event described in the matrix clause. For instance, on a natural interpretation, the RC in (6) not only restricts the reference of the kindergartner, but also carries an implicature that the student’s tardiness is the reason for the reproach.

(6) The teacher reproached the kindergartner who always shows up late.

This implicature depends on the world knowledge that persistent lateness is a plausible reason to be reproached, and thus, crucially, this implicature is not triggered by a property internal to the RC itself. That is, the same RC in a different context can merely serve to modify the preceding NP if the world knowledge necessary to support a causal interpretation does not exist, as in (7).

(7) The teenager babysits the kindergartner who always shows up late.

Likewise, it is possible to use a different RC with the matrix clause of (6) such that a causal relationship is not inferred:

2The fact that this is an implicature and not an entailment is demonstrated by the fact that it is cancelable, as in (i).

(i) The teacher reproached the kindergartner who always shows up late.

The kindergartner had forgotten his lunch money for the third day in a row.
In (8), for instance, the comprehender is not normally led to the inference that reminding the speaker of Harry Potter is the reason for the reproach.

This kind of intrasentential inference process mirrors the process that comprehenders use to establish discourse coherence across sentences, as already described with respect to (1) and (2). That is, the causal reasoning that links the matrix and relative clauses in (6) is similar to that used in establishing an Explanation relationship between the two sentences in (9). Similarly, the lack of inferred causation in (7) mirrors the non-causal Elaboration relation that holds between the sentences in (10).

Paralleling our observations regarding (6) and (7), the different relationship between the sentences in (9) and (10) stems from the fact that lateness provides a plausible reason to reproach someone but not a plausible reason to babysit someone.

1.5. Observation 3: Implicit Causality Influences Next Mention Expectations

Our third observation is already well established in the literature: That IC verbs impute causality primarily to one of the participants of the eventuality they denote, and thus create a strong bias toward mentioning that participant in any ensuing explanation of that eventuality (Garvey and Caramazza, 1974; Brown and Fish, 1983; Au, 1986; McKoon et al., 1993). Some verbs, like \textit{detest} in (4) and (11), are OBJECT-BIASED, meaning that it is the direct object that comprehenders expect to hear mentioned again in connection with an explanation: If John detests Mary, then the cause is likely to originate from a property of Mary. On the other hand, verbs like \textit{annoy} in (12) are SUBJECT-BIASED: If John annoys Mary, then the cause presumably originates from a property of John. Non-IC verbs, such as \textit{babysit} in (13), are reported to have weaker and less consistent biases.

(8) The teacher reproached the kindergartner who reminds me of Harry Potter.

(9) The teacher reproached the kindergartner. The kindergartner always shows up late.

(10) The teenager babysits the kindergartner. The kindergartner always shows up late.

(11) John detests Mary because... [OBJECT-BIASED IC VERB] ...she is rude and arrogant.

(12) John annoys Mary because... [SUBJECT-BIASED IC VERB]
...he is rude and arrogant.

(13) John babysits Mary because... [NON-IC VERB]...he needs the money. / ...she is too young to be left alone.

These next-mention biases have commonly been measured using story completion tasks with because-prompts like those in (11)-(13) (typical continuations are listed below each prompt). Kehler et al. (2008) also found that a similar pattern of biases emerges for prompts like (11)-(13) even without the because connective when, crucially, only the subset of continuations in which an Explanation relation is operative are analyzed. That is, the bias is tied specifically to causes per Explanation relations, and is not dependent on the conjunction because.

1.6. Implicit Causality and RC Attachment

We are now ready to introduce the manipulation of interest, illustrated in the matrix clauses of (14)-(15), with sample RCs shown below in (a-b):

(14) John babysits the children of the musician who ...
    a. ...lives in La Jolla. [low]
    b. ...are students at a private school. [high]
(15) John detests the children of the musician who...
    a. ...lives in La Jolla. [low]
    b. ...are arrogant and rude. [high]

The matrix clauses in these examples differ only in the verb: detests is an object-biased IC verb, whereas babysits is non-IC. The default low-attachment preference attested in English predicts uniform biases across (14)-(15): for instance, in a passage completion experiment, we would expect to see more low-attaching completions (like (14a) and (15a)) than high-attaching ones (like (14b) and (15b)). We would likewise expect the RC verb lives in (14a) and (15a), which agrees in number with the lower NP, to be easier to process on-line than the verb are in (14b) and (15b), which agrees with the higher NP. However, if comprehenders utilize the coherence-based pragmatic information described in Sections 1.3-1.5, we might expect to see a difference in these measures: The bias toward high attachments should be greater in (15) than in (14). That is, if IC verbs create a strong expectation that an explanation will follow, and such an explanation is likely to re-mention the direct object of the IC verb – which, crucially, would be the high attachment point for an ensuing RC – then we would expect an RC that provides an explanation to attach high.
This reasoning can be illustrated more concretely with (and indeed, can be seen as a natural consequence of) a rational probabilistic model such as that described in equation I.

\[
p(attachRC_{\text{high}}|\text{verb}) = \\
p(mentionNP_{\text{high}}|exp, \text{verb}) * p(exp|\text{verb}) + \\
p(mentionNP_{\text{high}}|\text{non-exp, verb}) * p(\text{non-exp}|\text{verb})
\]

This equation captures the fact that the attachment height of an RC is conditioned on the matrix clause verb \(p(attachRC_{\text{high}}|\text{verb})\) and can be modeled with information about the probability of particular upcoming coherence relations \(p(exp|\text{verb}), p(\text{non-exp}|\text{verb})\) and the probability of re-mentioning the direct object given the operative coherence relation \(p(mentionNP_{\text{high}}|exp, \text{verb}), (p(mentionNP_{\text{high}}|\text{non-exp, verb})\)). We will step through the terms of this equation below explaining their relation to the observations in the previous subsections. Note that in these contexts the choice to re-mention the direct object corresponds to a choice to attach an RC to the NP in the high-attachment position.

We saw in Section 1.3 that clauses containing IC verbs, such as \textit{detest} in (15), create a strong expectation for an ensuing explanation (i.e., \(p(exp|\text{verb}_{\text{IC}})\) is relatively high). We then saw in Section 1.4 that this explanation can be delivered via an immediately-following RC. Finally, we saw in Section 1.5 that if an explanation were to follow, an object-biased verb like \textit{detest} creates a strong expectation that the explanation will re-mention the verb’s direct object (i.e., \(p(mentionNP_{\text{high}}|exp, \text{verb}_{\text{IC}})\) is relatively high). Whereas we would expect this same system of coherence-based knowledge to exert parallel influences in non-IC examples like (14), the biases toward high attachment are expected to be considerably weaker, since non-IC verbs do not create a strong expectation toward an ensuing explanation \(p(exp|\text{verb}_{\text{non-IC}})\) is comparatively low as per Section 1.3), nor is the direct object favored to be the locus of such explanations \(p(mentionNP_{\text{high}}|exp, \text{verb}_{\text{non-IC}})\) is comparatively low as per Section 1.5). We likewise expect \(p(mentionNP_{\text{high}}|\text{non-exp, verb})\) to be comparatively low for both verb types. As such, we conjecture that upon hearing matrix sentences containing IC verbs like (15), comprehenders’ pragmatic knowledge will yield an expectation that the upcoming RC will provide an explanation that expresses a predication about (and therefore attaches high to) the direct object significantly more often than for matrix sentences containing non-IC verbs as in (14).

Crucially, this reasoning only goes through by making a significant assumption: That comprehenders, having processed the initial part of a matrix clause, implicitly integrate all three types of pragmatic knowledge and biases discussed in Sections 1.3-1.5 and use them when making a \textit{syntactic} attachment decision.
mid-sentence. This stimulus design therefore affords an unprecedented opportunity to examine a set of strong predictions concerning the potential effects of pragmatic biases on incremental syntactic analysis. We submit, therefore, that a difference in the comprehender’s behavior in processing passages like (14) and (15) in the predicted direction would constitute a novel and significant piece of evidence in support of an expectation-based model of syntactic processing that incorporates a broad set of cross-modular information sources.

The remainder of this paper presents two experiments that test this hypothesis. If comprehenders are indeed using coherence-driven biases mid-sentence, then one would expect to see effects with respect to the types of RC completions they generate following IC matrix-clause verbs (Experiment 1, a sentence completion study). If these biases contribute to on-line processing, then one would expect to see processing difficulty associated with those RC attachments that violate the biases introduced by the matrix-clause verb (Experiment 2, a self-paced reading study).

2. Experiment 1: Sentence Completion Study

The first experiment uses an off-line sentence-completion task to test the hypothesis that object-biased IC verbs influence sentence completions for ambiguously-attached RCs in ways consistent with the three observations above. The prompt pair in (16) involves a single manipulation (henceforth referred to as ‘verbtype’)—whether the main clause has an IC verb (16b) or a non-IC verb (16a).

(16) a. John babysits the children of the musician who ...

b. John detests the children of the musician who ...

According to our hypothesis, comprehenders will utilize discourse-level expectations about upcoming coherence relations mid-sentence. In that case, the presence of an IC verb in (16b) should elicit a greater number of completions that explain the matrix clause event than elicited by the non-IC variant (16a), and, in light of the fact that detests is object-biased, these explanations should tend to attach high. If, on the other hand, such expectations are not utilized during syntactic processing, RCs following both verbtypes should attach low.

2.1. Methodology

Participants
Fifty-two monolingual English-speaking undergraduates at UC San Diego participated in the study for course credit in Linguistics courses.

Materials

Stimuli consisted of twenty-one pairs of sentences differing only in the matrix verb, as in (16); the complete stimuli can be found in the appendix. The subject of the matrix verb in each sentence was a proper name, and the direct object consisted of a complex NP containing two NPs connected by the genitive marker of. Both NPs denoted human referents so that the relative pronoun who could plausibly be used to modify either NP. In order to make disambiguation easier for the judges, the complex NP consisted of a singular NP and a plural NP so that number agreement on the embedded verb could be used to assess the intended attachment site of the RC. The order of singular and plural was balanced across stimuli (10 singular-plural, 11 plural-singular).

The verb in the matrix clause was either an object-biased IC verb or a non-IC verb. IC verbs were selected from two lexical semantic categories that Levin (1993) labels ‘psych’ and ‘judgment’ verbs. Non-IC verbs were adapted from those used by McKoon et al. (1993) in their study of IC and pronoun interpretation. For our stimuli, psych verbs appeared in the present tense since they describe non-eventive states (e.g., detest, adore), whereas judgment verbs appeared in the simple past (e.g., scolded, praised). Each pair of IC and non-IC verbs was matched for tense as in (16).

In addition to the experimental items, the experiment included twenty-one fillers and twenty-one additional stimuli for an unrelated experiment, pseudorandomized to create eight lists. The additional fillers consisted of sentences with non-transfer verbs and a variety of prompts as well as sentences with complex NPs and unambiguous RC prompts. Half of the unambiguous RC fillers enforced a low attachment and half enforced a high attachment.

Procedure

Sentence completions were collected via a web-based interface that participants could access from their own computer. Each item was presented on a page by itself with a text box in which participants were instructed to write their completion. The entire experiment took roughly thirty minutes, but participants were encouraged to have an hour available so that the experiment could be com-

---

3 The stimuli for the interleaved experiment contained sentences with transfer-of-possession verbs followed either by a full stop and a completion prompt or a full stop and an ambiguous pronoun prompt: Matt passed a sandwich to David. (He) ...

4 The unambiguous RC prompts used the relative pronoun who in contexts in which only one of the two nouns in the complex NP was human.
pleted in one session. (Participants could leave and return at a later time by identifying themselves with an ID number.) They were instructed to imagine a natural sentence completion for each prompt, writing the first completion that came to mind and avoiding humor.

**Evaluation and Analysis**

Two trained judges—the first author of this paper and an undergraduate Linguistics student—annotated all responses for the relationship that could be inferred to hold between the RC and the matrix clause (‘explanation’ and ‘non-explanation’) and the RC’s intended attachment site (low or high). An RC was labeled ‘non-explanation’ if it provided additional information about one of the nouns without providing additional causal information about the event in the matrix clause *per se*. Categorization as an explanation RC, on the other hand, required that a causal link be inferrable between the information conveyed by the matrix clause and the information in the RC. Disagreements were resolved through discussion. The intended attachment site was assessed in light of the matrix clause context and the elicited RC. An RC was excluded from the analysis if at least one coder assessed its attachment height to be ambiguous, as well as in the few cases in which the coders disagreed.

The sample completions in (17) and (18) illustrate explanation and non-explanation RCs that attach either low or high. (Examples of actual completions corresponding to each of (17a)-(17d) and (18a)-(18d) can be found in the appendix.) Note that the sample non-explanation RCs listed do not differ between the non-IC and IC contexts ((17a)-(17b), (18a)-(18b)) because such RCs need not reflect information about the event described in the matrix clause. Explanation RCs, on the other hand, provide an explanation of the matrix clause event and therefore are shown varying with the matrix clause ((17c)-(17d), (18c)-(18d)).

(17) **Non-IC verb**: John babysits the children of the musician who ...

- a. ...lives in La Jolla.  
- b. ...are in elementary school.  
- c. ...works a late shift every night.  
- d. ...are left home on Friday nights.

(18) **IC verb**: John detests the children of the musician who ...

- a. ...lives in La Jolla.  
- b. ...are in elementary school.  
- c. ...encourages their 3am drum solos.  
- d. ...are arrogant and rude.
As (17) and (18) show, both verb types can be felicitously followed by an explanation or a non-explanation RC, and neither RC type enforces a particular attachment level. The hypothesis is that the combination of coherence biases and next-mention biases will render high-attaching explanation RCs more likely following IC verbs than non-IC verbs: Completions like (18d) will be more common than (17d). The low-attaching non-explanation RCs, on the other hand, are predicted to be more expected following non-IC verbs than IC verbs: Completions like (17a) will be more common than (18a).

We conducted analyses of variance on the assessed RC completion types and on the assessed attachment sites to test for a main effect of verbtype. Because these measures involve examining proportions of binary categorical outcomes, we first applied an arcsine transformation (Sokal and Rohlf, 1995) to subject- and item-specific percentages in each condition. For clarity, we present grand means as back-transformed percentages. The observed RC types and RC attachment sites were also modeled using mixed-effects multinomial logistic regressions with random subject-specific and item-specific intercepts (Jaeger, 2008). We report the coefficient estimate and p-value (based on the Wald Z statistic; Agresti, 2002) for the factor verbtype in models fitted to the observed RC completion types and to the observed RC attachments.

2.2. Results

As predicted, IC verbs yielded significantly more explanation-providing RCs than non-IC verbs (main effect of verbtype on RC type: $F_1(1,51)=292.22$, $p<0.001$; $F_2(1,20)=87.665$, $p<0.001$) and significantly more high-attaching RCs (main effect of verbtype on attachment height: $F_1(1,51)=27.158$, $p<0.001$; $F_2(1,20)=6.8475$, $p<0.05$). In the logistic regressions, verbtype was a significant factor for modeling the binary outcome of RC type (whether an RC provided an explanation: $\beta=4.530$, $p<0.001$) and the binary outcome of attachment height (whether the RC attached high: $\beta=0.803$, $p<0.005$).

As can be seen in Figure 1a, verbtype affected the types of RCs participants produced: More than half (63.9%) of the RCs following IC verbs provided an explanation of the event in the matrix clause, while only a small proportion

---

5This analysis represents a conservative analysis in which an RC was excluded if at least one coder assessed it as ambiguous (22.5% of the total number of RC completions). The results remain significant if RCs are included if at least one coder assigned a non-ambiguous interpretation (Attachment: $F_1(1,51)=53.52$, $p<0.001$; $F_2(1,20)=8.1197$, $p<0.01$; RC type: $F_1(1,51)=356.07$, $p<0.001$; $F_2(1,20)=96.407$, $p<0.001$). We restricted the analysis to subject-extracted RCs since object-extracted RCs made up fewer than 1% of the total completions, and their inclusion does not affect the overall results.
(11.0%) of RCs following non-IC verbs provided an explanation. Figure 1b shows that the pattern of RC attachment differs by verb type as well. In the non-IC context, only 36.5% of the unambiguous elicited completions contained a high attachment, which matches the reported low-attachment preference for English. In the IC context, the low-attachment preference disappears with 50.6% of unambiguous completions containing a high attachment. All of these proportions represent subject means and are shown in Figure 1 with error bars for standard error of the mean.

Figure 2 shows the results broken down by verb type and RC type. Regardless of verb type, explanation-providing RCs had a higher incidence of high attachment (66.3% for IC verbs, 47.0% for non-IC verbs) than RCs that did not provide explanations (26.0% for IC verbs, 35.9% for non-IC verbs). Pairwise comparisons of explanation vs. non-explanation RCs were significant in the IC condition ($F_1(1,49)=35.351, p<0.001$; $F_2(1,20)=36.419, p<0.001$; logistic regression: $\beta=2.9391, p<0.001$) but not significant in the non-IC condition ($F_1(1,32)=0.4819, p=0.49$; $F_2(1,8)=0.6325, p=0.45$; logistic regression: $\beta=0.6246, p=0.15$). The greater incidence of high attachment for explanation RCs in the IC condition follows intuitively from the fact that the IC verbs used in this experiment impute causality to their direct object (the high NP). Even for non-IC verbs, it is unsurprising that explanation RCs might tend to attach to the high NP more than non-explanation RCs, since the high NP is a direct participant in the matrix-clause event being explained, whereas the low NP is not (though see Section 2.3). Unlike the case for IC verbs, however, the difference in the proportion of high attachments was not significant.

2.3. Discussion

The sentence completion study was designed to test the hypothesis that an ensemble of pragmatic factors would conspire to yield more explanation-providing
RCs — and as a result, a higher percentage of high-attaching RCs — in sentences with IC verbs than in sentences with non-IC verbs. Both of these predictions were confirmed. The pattern of high attachments in IC contexts therefore shows that models that predict a uniform low-attachment preference are inadequate. These results instead suggest that a variety of factors determine attachment biases; whereas in many contexts these factors conspire to yield a low attachment bias, in others they yield a high attachment bias. The fact that RC attachment biases proved to be dependent on the discourse relation between the RC and the matrix clause suggests that the relevant factors go beyond merely those derivable from properties of the NPs themselves; in particular, coherence-driven factors determine attachment biases as well.

The results also confirm Kehler et al.’s (2008) conclusion that comprehenders are sensitive to two types of biases invoked by IC verbs: a clause-level coherence bias toward upcoming explanations and an entity-level next-mention bias conditioned on the presence of an explanation relation. Whereas Kehler et al.’s study showed the effect of IC biases on intersentential pronoun interpretation, the current findings go further in demonstrating that explanations can be inferred to hold intrasententially so as to affect a syntactic interpretation decision. As such, the linking together of elements of a discourse into a larger coherent structure cannot be cast as a process that occurs only after sentences have been digested in their entirety. Indeed, not only are coherence relations anticipated mid-sentence, but coherence-driven expectations affect the process of building the very syntactic structure that will yield the meanings over which coherence will ultimately be established.

When we consider the two types of biases invoked by IC verbs, our results further suggest that the clause-level bias towards an upcoming explanation plays a stronger role in the attachment decision than the next-mention bias toward the
direct object. Recall that according to equation I, repeated below in II, the way in which the preceding context influences attachment preferences is complex, depending simultaneously on biases about upcoming coherence relations and biases about next-mention given the operative coherence relation and the presence of an object-biased IC verb.

\[
p(attachRC_{\text{high}}|\text{verb}) = \frac{p(mentionNP_{\text{high}}|exp, verb)}{p(mentionNP_{\text{high}}|\text{verb})} \times p(exp|\text{verb}) + \frac{p(mentionNP_{\text{high}}|\text{non-exp}, verb)}{p(mentionNP_{\text{high}}|\text{verb})} \times p(\text{non-exp}|\text{verb})
\]  

(II)

Overall, the results from Experiment 1 confirmed that the bias to an upcoming explanation is strongly conditioned on verbtype (i.e., \(p(exp|\text{verb}_{IC})\) was much larger than \(p(exp|\text{verb}_{non-IC})\), as per Figure 1a) and furthermore that the bias to re-mention the higher NP is conditioned on the combined information regarding verbtype and the inference of an explanation relation (i.e., \(p(mentionNP_{\text{high}}|\text{exp}, \text{verb}_{IC})\) was larger than both \(p(mentionNP_{\text{high}}|\text{non-exp}, \text{verb}_{IC})\) and \(p(mentionNP_{\text{high}}|\text{exp}, \text{verb}_{non-IC})\), as per Figure 2). Of these two types of biases, however, it is the coherence bias that most clearly differentiates the IC and nonIC conditions. While participants were slightly more likely to favor the higher NP in explanation-providing RCs in the IC condition than in the nonIC condition, the overall pattern of attachment is driven by the strong bias to produce more explanations in the IC condition than in the non-IC condition, such that the difference between \(p(\text{coh}|\text{verb}_{IC})\) and \(p(\text{coh}|\text{verb}_{non-IC})\) more strongly dictates the attachment outcome, as compared to \(p(mentionNP_{\text{high}}|\text{coh}, \text{verb}_{IC})\) and \(p(mentionNP_{\text{high}}|\text{coh}, \text{verb}_{non-IC})\)

Although the experiment confirmed our predictions regarding the IC biases, a closer analysis of the elicited sentence completions revealed that several factors likely conspired to reduce the effect from what might otherwise have been found. In particular, our analysis identified two subpatterns of behavior that are not apparent in the aggregate effects. The first subpattern concerns the coherence bias, where we found that some verbs that have been classified in the literature as non-IC actually yielded a larger number of explanation-providing RCs than some IC verbs — that is, \(p(exp|\text{verb}_{IC})\) was not always larger than \(p(exp|\text{verb}_{non-IC})\). For example, the verb watch, which McKoon et al. (1993) included in a non-IC condition, yielded 46.2% explanations — more than some IC verbs such as like (26.7% explanations) and value (22.7% explanations). The appendix contains the full list of the percentages of explanation RCs that each verb elicited. Caramazza et al. (1977) previously commented that the next-mention biases of IC verbs lie along a continuum; here we find that the same is true for their biases towards yielding ensuing explanations as well. We therefore would have expected a stronger effect if the IC verbs used had uniformly stronger biases towards explanations than
their non-IC counterparts.

The second subpattern concerns the attachment biases, where we found that the presence of both an IC verb and an explanation-providing RC are not in themselves sufficient to yield a high-attachment bias — that is, \( p(\text{mentionNP}_{\text{high}}|\text{exp}, \text{verb}_{\text{IC}}) \) is not always larger than \( p(\text{mentionNP}_{\text{low}}|\text{exp}, \text{verb}_{\text{IC}}) \). The relation between the NPs in the complex NP also has an influence. The two IC items in (19) show how specific complex NPs can shift the bias to yield more high-attaching or more low-attaching RCs.

\[
\begin{align*}
(19) & \quad \text{a. Alan punished the accountant of the businessmen who ...} \\
& \quad \text{b. Bill congratulated the teacher of the second graders who ...}
\end{align*}
\]

Example (19a) yielded a large proportion of explanation-providing RCs (85.7%), and those RCs consistently attached to the higher NP (100%). Example (19b) also yielded many explanation-providing RCs (81%), but in this case, the RCs tended to attach low (only 29% high attachment). Example (19b) differs from (19a) in that the lower NP in (19b) refers to a set of individuals (NP2: the second graders) who are under the care or responsibility of the individual referred to by the direct object (NP1: the teacher). Because of the possibility of attributing responsibility to the NP1 referent for the NP2 referent’s behavior, the explanation-providing RCs could plausibly mention either the guardian or the guardian’s wards.\(^6\) As such, avoiding such relationships between NP1 and NP2 would presumably also have yielded a stronger effect.

Although Experiment 1 confirmed our predictions, the results are restricted to an off-line completion task. If, as we hypothesize, the coherence-driven biases that emerge are indeed deployed mid-sentence, then one would expect to see effects in comprehenders’ incremental processing in a self-paced reading time experiment. The goal of Experiment 2 is to test this hypothesis.

3. Experiment 2: Self-Paced Reading Study

Experiment 2 was designed to test whether the attachment biases revealed by Experiment 1 generate expectations that are deployed rapidly in on-line comprehension. Crucially, we examine the prediction that inferences about intraclausal coherence relations are fully incremental, and can therefore affect local syntactic disambiguation before comprehenders have been exposed to complete clauses.

\(^6\)Similarly behaving items included scold the landlady of the actors who..., detest the father of the students who..., and pity the bodyguards of the celebrity who...
To examine this question, we conducted a moving-window self-paced reading study. We adapted the stimuli from Experiment 1 to create a 2 × 2 design, varying the verb type and the RC attachment height, as in (20) and (21). Underscores connect words presented together as a single region in the study.

(20) **NON-IC MATRIX:** John babysits the children of the musician ...
   a. [LOW ATTACHMENT] ... who is generally arrogant and rude.
   b. [HIGH ATTACHMENT] ... who are generally arrogant and rude.

(21) **IC MATRIX:** John detests the children of the musician ...
   a. [LOW ATTACHMENT] ... who is generally arrogant and rude.
   b. [HIGH ATTACHMENT] ... who are generally arrogant and rude.

The attachment height of the RC in (20) and (21) is signaled by the number agreement information associated with the finite verb (e.g., the verb *is* in (20a) agrees in number with NP2, the NP at the low-attachment site). If the use of the pragmatic biases detailed in Sections 1.3 through 1.5 is fully incremental, the comprehender should already have formed expectations about the attachment height of the RC initiated by *who* before encountering the finite verb. The degree to which these expectations match the RC attachment height signaled by the finite verb should then determine the difficulty of processing that verb. The finite verb therefore constitutes the critical region of the experiment. Because differences in processing difficulty in self-paced reading often show up a region or two downstream of the critical region (see Mitchell, 1984), especially when the critical region is short as it is here, the immediately post-critical word (*generally* in (20) and (21)) was always an adverb chosen to be relatively non-indicative of attachment height. This word and the immediately following word (*arrogant* in (20) and (21)) are the spillover regions for this experiment. On the hypothesis that coherence biases affect the online formation of attachment preferences, we expect an interaction between attachment height and verb type on reading times at the critical and/or spillover regions, where we predict a low-attachment bias to emerge for non-IC verbs ((20a) should be easier than (20b)) and a reduced, neutralized, or even reversed bias to emerge following IC verbs ((21a) should be not as much easier than, equally difficult as, or even harder than (21b)). On the other hand, if our hypothesis is incorrect — that is, integration into the larger discourse structure occurs only after the entire sentence has been processed — then the low-attachment bias should hold across the board ((20a),(21a) should be easier than (20b),(21b)) in the critical and/or spillover regions. As such, the attachment height × verb type interaction is not predicted by either default-bias or clausal-integration accounts because these accounts fail to provide a mechanism by which pragmatic biases of the matrix clause verb could influence attachment
height and do so mid-sentence.

3.1. Methodology

We observed in Section 2.3 that for certain stimuli used in Experiment 1, the relationship between the high and low NPs was such that an explanation relation remained strongly compatible with reference to the low NP. Because our analysis and the default low-attachment account make the same predictions in such cases, and hence they cannot be used to discriminate between the analyses, we sought to avoid such stimuli for this experiment. Similarly, we also observed in Section 2.3 that the biases towards ensuing explanations for the non-IC and IC verbs were not as consistent across verbs as expected, which also affected the extent to which the two accounts could be discriminated. We therefore selected IC verbs from Experiment 1 that had strong biases towards explanation RCs and supplemented them with a few additional verbs taken from Kehler et al.’s (2008) study (which were in turn taken from McKoon et al.’s (1993) study, with some minor substitutions). The 20 non-IC verbs similarly consisted of a mix of verbs from McKoon et al. (1993), Levin (1993), and from Experiment 1.

Participants

58 monolingual English speakers participated in the reading time experiment for credit in Linguistics and Psychology courses.

Materials

Each experimental item consisted of a matrix clause with a proper name, a verb, and a complex NP direct object, followed by a temporarily ambiguous RC, as in (20) and (21). The complex NP contained a singular NP and a plural NP so that number agreement on the critical embedded verb (is/are in (20) and (21)) served to disambiguate the attachment site of the RC. The order of singular and plural in the complex NP was balanced across stimuli so that high attachment was signaled with plural agreement for half the items and with singular agreement for the other half. The embedded verb (from here on, the ‘RC verb’) was always a be or have verb form that was inflected for number agreement; depending on the item, it served either as an auxiliary or as the main verb of the relative clause. A semantically neutral adverb always appeared immediately after the critical region. The complete set of experimental items can be found in the appendix.

The experiment consisted of 10 practice items, followed by 20 experimental items mixed with 30 fillers, pseudorandomized for each subject. The filler items were similar to the stimuli in that some included proper names and RCs or other subordinate clauses.

Procedure
Items were presented in a moving-window self-paced reading paradigm, using DMDX experiment software (Forster and Forster, 2003). Sentences appeared in white letters on a dark background, left-justified on a 19” CRT screen, and no sentence was longer than one line of text. Sentences initially appeared as a series of dashes (— — — —) obscuring the words, and participants pushed a button on a Logitech USB gamepad to reveal each region. The presentation was non-cumulative such that previous regions were replaced with dashes when the next region appeared. The critical region and the spillover regions were revealed one word at a time, but multi-word regions were used elsewhere to present short phrases such as a verb and a preposition (*stared at, stood near*) or a determiner and a noun (*the children*). Multi-word regions are indicated in the stimuli set in the appendix. Participants pushed either a YES or NO button on the gamepad to answer a comprehension question after every sentence, and they received automatic feedback whenever they answered incorrectly. They were instructed to read as quickly and carefully as possible, making sure they understood the complete sentence and slowing down if they answered multiple questions incorrectly. We recorded reading times for each region as well as the participant’s response to the comprehension question.

3.2. Results

After excluding three participants whose comprehension-question accuracy was not significantly better than chance, the percentage of correct responses was 93.03% for fillers and 85.07% for experimental items (percentages over subject means), indicating that participants paid attention to the task. Table 1 shows the raw reading times by condition for the critical region and the spillover regions, as well as the mean accuracy on comprehension questions for each condition.

<table>
<thead>
<tr>
<th></th>
<th>RC Verb</th>
<th>Spillover1</th>
<th>Spillover2</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>IC.high</td>
<td>395.70 ±16.83</td>
<td>430.43 ±18.90</td>
<td>442.81 ±18.84</td>
<td>.873 ±.02</td>
</tr>
<tr>
<td>IC.low</td>
<td>398.83 ±16.71</td>
<td>474.16 ±23.26</td>
<td>477.19 ±26.34</td>
<td>.780 ±.02</td>
</tr>
<tr>
<td>nonIC.high</td>
<td>402.03 ±16.55</td>
<td>501.48 ±24.26</td>
<td>473.59 ±20.22</td>
<td>.862 ±.02</td>
</tr>
<tr>
<td>nonIC.low</td>
<td>403.96 ±13.83</td>
<td>462.63 ±20.03</td>
<td>437.50 ±15.91</td>
<td>.887 ±.02</td>
</tr>
</tbody>
</table>

Table 1: Raw RTs and question accuracy (subject means ± standard error)

Figure 3 shows the residual reading times for each of the four conditions starting at the matrix verb. Comprehension-question accuracy and reading times were analyzed with 2×2 ANOVAs, by subject and by item. The results were also analyzed using linear mixed-effects models with random subject-specific and
item-specific intercepts. For these models, we report the coefficient estimates and MCMC-derived p-values (Baayen et al., 2008).

Comprehension question accuracy

All question-accuracy ANOVAs were conducted on arcsine transformed proportions of correct answers. There was a marginal main effect of attachment height favoring high-attaching RCs (marginal by subject: $F_{1}(1,54)=3.889$, $p=0.054$; $F_{2}(1,19)=2.778$, $p=0.112$). There was also a main effect of verb-type favoring non-IC verbs (significant only by subject: $F_{1}(1,54)=4.59$, $p<0.05$; $F_{2}(1,19)=2.206$, $p=0.154$). These main effects are driven by an interaction in which low-attaching RCs in the IC condition yielded lower accuracies than any other condition. The interaction that emerges is consistent with the predicted interaction for processing difficulty in cases in which the RC violates the expectations generated from the preceding context (verbtype×attachment-height interaction significant by subject and marginal by item: $F_{1}(1,54)=7.346$, $p<0.01$; $F_{2}(1,19)=3.89$, $p=0.063$). In a mixed-effects logistic regression, attachment height was not a significant factor for modeling question accuracy (attachment: $\beta=-0.273$, $p=0.13$), whereas verbtype and the verbtype×attachment interaction were significant (verbtype: $\beta=0.445$ $p<0.05$; verbtype×attachment interaction: $\beta=0.972$, $p<0.01$).

Reading time results
We analyzed residual reading times at the critical region and two spillover regions. Residual reading times adjust for overall differences in participants’ reading rates as well as differences in readers’ sensitivity to word length. Residual RTs were calculated as the difference between the actual reading time on a word and the reading time predicted by a regression equation (computed separately for each participant, using all experimental and filler items) relating word length to reading time (Trueswell et al., 1994). We removed residual RTs that were more than four standard deviations away from the mean, per region and per condition (0.36% of the data). The following results represent an analysis of all non-outlier items, regardless of comprehension-question accuracy.

At the disambiguating RC verb (is/are), there were no significant effects for verbtype ($F_{1}(1,54)=1.071$, $p=0.31$; $F_{2}(1,19)=1.124$, $p=0.30$), or the verbtype×attachment interaction ($F_{1}<1$). In a mixed-effects linear regression, the factors for verbtype, attachment height, and the verbtype×attachment interaction were not significant factors for modeling residual reading time (verbtype: $\beta=-0.1556$, $p=0.99$; attachment: $\beta=9.84$, $p=0.31$; verbtype×attachment interaction: $\beta=-13.258$, $p=0.51$).

At the first spillover region (generally), there were again no main effects of verbtype ($F_{1}<1$) or attachment height ($F_{1}(1,54)=1.295$, $p=0.26$, $F_{2}<1$). However, a significant interaction was observed in the predicted direction: High attachments were read more slowly than low attachments in non-IC conditions but faster in IC conditions ($F_{1}(1,54)=5.522$, $p<0.05$; $F_{2}(1,19)=6.167$, $p<0.05$). In a mixed-effects linear regression, the factors for verbtype and attachment were not significant (verbtype: $\beta=8.672$, $p=0.47$; attachment: $\beta=12.027$, $p=0.31$), whereas the verbtype×attachment interaction was (verbtype×attachment interaction: $\beta=-63.60$, $p<0.01$). To test for full crossover interaction, we conducted pairwise tests of high vs. low attachment separately in IC versus non-IC conditions. In the IC condition there was an effect of attachment level marginal by subjects, significant by items, and significant in a mixed-effects model ($F_{1}(1,54)=3.45$, $p=0.069$; $F_{2}(1,19)=7.91$, $p<0.05$; $p_{MCMC}<0.05$). In the non-IC condition the effect was insignificant ($F_{1}(1,54)<1$; $F_{2}(1,19)=1.09$, $p=0.31$; $p_{MCMC}=0.12$).

At the second spillover region (arrogant), there were again no main effects ($F_{1}<1$), but the same interaction was significant ($F_{1}(1,54)=6.588$, $p<0.05$; $F_{2}(1,19)=4.967$, $p<0.05$). In a mixed-effects linear regression, the main factors of verbtype and attachment height were not significant, but the interaction was (verbtype: $\beta=3.106$, $p=0.78$; attachment: $\beta=-3.279$, $p=0.77$; verbtype×attachment interaction: $\beta=-47.10$, $p<0.05$). To test for full crossover interaction, we conducted pairwise tests of high vs. low attachment separately in IC versus non-IC conditions. In the IC condition the effect was not sig-
significant ($F_1(1,54)=2.51$, $p=0.12$; $F_2(1,19)=1.97$, $p=0.18$; $p_{MCMC}=0.12$). In the non-IC condition the effect was significant by subjects and by items, and marginal in a mixed-effects model ($F_1(1,54)=5.8$, $p<0.05$; $F_2(1,19)=4.37$, $p<0.05$; $p_{MCMC}=0.08$).

We also conducted an analysis of RTs summed across the two spillover regions, since a significant interaction pattern was observed in both regions. We found no main effects ($F$s<1) and an interaction significant by both subjects and items ($F_1(1,54)=10.05$, $p<0.01$; $F_2=9.56$, $p<0.01$; $p_{MCMC}<0.001$). To test for full crossover interaction, we conducted pairwise tests of high vs. low attachment separately in IC versus non-IC conditions. In the IC condition the effect was significant ($F_1(1,54)=6.79$, $p<0.05$; $F_2(1,19)=9.08$, $p<0.01$; $p_{MCMC}<0.005$). In the non-IC condition the effect was marginal by subjects and significant in the mixed model analysis ($F_1(1,54)=3.77$, $p=0.058$; $F_2(1,19)=2.33$, $p=0.14$; $p_{MCMC}<0.05$).

Analyses of the raw reading times were qualitatively the same, as were analyses of residual RTs with incorrectly answered items excluded.\footnote{Considering the raw reading times, there were no effects at the disambiguating verb (F$s<1$); regression with factors for verbtype: $\beta=4.112$, $p=0.69$; attachment: $\beta=5.577$, $p=0.61$; verbtype$\times$attachment: $\beta=-7.953$, $p=0.72)$. At the first spillover region, there was an effect of verbtype by subject and a significant interaction (verbtype: $F_1(1,54)=7.075$, $p<0.05$; $F_2(1,19)=3.548$, $p=0.075$; attachment: $Fs<1$; verbtype$\times$attachment: $F_1(1,54)=6.853$, $p<0.05$; $F_2(1,19)=5.434$, $p=0.05$; regression for spillover1 with factors for verbtype: $\beta=30.466$, $p<0.05$; attachment: $\beta=1.787$, $p=0.88$; verbtype$\times$attachment: $\beta=-81.31$, $p<0.005$). At the second spillover region, there were no main effects (F$s<1$) and the same interaction was significant ($F_1(1,54)=6.075$, $p<0.05$; $F_2(1,19)=6.078$, $p<0.05$; regression for spillover2 with factors for verbtype: $\beta=-2.968$, $p=0.83$; attachment: $\beta=-2.738$, $p=0.81$; verbtype$\times$attachment: $\beta=-70.41$, $p<0.005$).

Considering the residual reading times with incorrectly answered items excluded, there were no effects at the disambiguating verb (verbtype: $F_1(1,54)=1.442$, $p=0.24$; $F_2(1,19)=2.116$, $p=0.16$; attachment: $Fs<1$; verbtype$\times$attachment: $Fs<1$; regression with factors for verbtype: $\beta=7.338$, $p=0.43$; attachment: $\beta=6.813$, $p=0.51$; verbtype$\times$attachment: $\beta=-12.77$, $p=0.53$). At the first and second spillover regions there were no main effects but significant interactions (Spillover1: verbtype: $F_1(1,54)=2.646$, $p=0.11$; $F_2(1,19)=1.701$, $p=0.21$; attachment: $Fs<1$; verbtype$\times$attachment: $F_1(1,54)=6.117$, $p<0.05$; $F_2(1,19)=5.216$, $p<0.05$; regression for spillover1 with factors for verbtype: $\beta=22.064$, $p=0.073$; attachment: $\beta=7.779$, $p=0.56$; verbtype$\times$attachment: $\beta=-74.05$, $p<0.005$; Spillover2: verbtype: $Fs<1$; attachment: $Fs<1$; verbtype$\times$attachment: $F_1(1,54)=7.598$, $p<0.01$; $F_2(1,19)=5.465$, $p<0.05$; regression for spillover2 with factors for verbtype: $\beta=0.1607$, $p=0.98$; attachment: $\beta=-2.4821$, $p=0.84$; verbtype$\times$attachment: $\beta=-64.89$, $p<0.01$).}
verbtype \times attachment: \beta = 29.613, p=0.40). No other pre-critical regions yielded significant effects.

3.3. An independent measure of verb bias

The foregoing results are supportive of our hypothesis that verbtype would affect initial RC attachment preferences and are consistent with the results of Experiment 1. In order to more thoroughly test the specific aspect of our hypothesis that the same preferences that are implicated in cross-sentence implicit-causality biases (Garvey and Caramazza, 1974; McKoon et al., 1993; Kehler et al., 2008) are also the key factors modulating online RC attachment preferences, we conducted an additional norming study on the main clauses of our experimental materials. Rather than writing story completions to fragments containing RC contexts as in Experiment 1, we asked participants to write a new sentence following a full stop as in (22) in order to establish an independent measure of the IC biases in these contexts:

(22) a. John babysits the children of the musician. ... [NON-IC]
  b. John detests the children of the musician. ... [IC]

By analyzing such data we can calculate (i) the proportion of continuations that explain the first sentence (as opposed to employing other coherence relations), and (ii) the proportion of explanation continuations in which the direct object is the next mentioned referent. These proportions may allow us to characterize IC bias as a gradient, rather than a categorical, characteristic of verbs and the contexts in which they are used. Two key predictions regarding these proportions follow from our hypothesis: first, in comparing the main clauses of the IC and non-IC conditions of our study, the IC condition is predicted to yield both a larger proportion of explanation continuations and a larger proportion of direct-object re-mentions in continuations in which an explanation coherence relation is operative; second, the resulting gradient measures of IC bias are expected to act as reliable predictors of attachment preferences (specifically, we should see a significant interaction between RC attachment and either of these independently observed measures of IC bias on reading times early in the RCs).\footnote{We are grateful to an anonymous reviewer for the suggestion to both norm our verbs and use the results in a regression analysis of RTs.}

It is important, however, to take into account one source of discrepancy between these cross-sentence next-mention biases and those in RC attachment as explored in Experiment 1: In contexts like (22), a greater-than-average bias toward re-mention of the direct object will come primarily at the expense of re-mentions
of the subject referent, whereas in the case of the RC contexts, attachments to the direct object will typically come at the expense of attachments to the referent in the object-of-PP position. Statistics for all three possible referents will therefore be used to test both of the key predictions described above.

3.3.1. Methodology

Forty monolingual English speakers wrote story completions following prompts like those in (22). Two trained judges—the first author of this paper and a Linguistics graduate student—assessed the 630 elicited completions that re-mentioned the subject NP (John), the direct object NP (the children), or the possessor NP (the musician). The judges annotated the responses for the completion type (‘explanation’ or ‘non-explanation’) and the choice of next mention (subject NP, direct object NP, or possessor NP).

3.3.2. Results: verbtype and gradient IC bias

Because these measures involve examining proportions of binary categorical outcomes, we conducted ANOVAs on arcsine-transformed subject- and item-specific means as we did for the completion study in Section 2. As in Section 2, we present means as back-transformed percentages and report results from mixed-effect multinomial logistic regressions.

The norming study confirmed that IC verbs yielded far more explanations (66.8%) than non-IC verbs (23.1%; main effect of verb type on completion type: $F_1(1,39)=65.83$, $p<0.001$; $F_2(1,19)=52.021$, $p<0.001$; regression: $\beta=2.081$, $p<0.001$). Column 1 of Table 2 shows the probability of an explanation completion (subject means) and the number of completions for each verbtype.

The norming study also confirmed that explanation completions following IC verbs contained the largest number of re-mentions of the direct object NP (67.2%), compared to explanation completions following non-IC verbs (8.6%), non-explanation completions following IC verbs (20.9%), and non-explanation completions following non-IC verbs (29.3%). In order to establish that the rate of object re-mentions is significantly larger in explanations following IC verbs

|          | $p(expl|\text{verb})$ | $p(\text{object} (\text{verb, expl})$ | $p(\text{object} | \text{verb, non-expl})$ |
|----------|----------------------|--------------------------------------|----------------------------|
| IC       | .668 (.335)          | .672 (.219)                         | .209 (.116)                |
| non-IC   | .231 (.295)          | .086 (.70)                          | .293 (.225)                |

Table 2: Probability of completion type (explanation vs. non-explanation) given verbtype and probability of object re-mention given verbtype and completion type. Support (n) in parentheses.
than in any other context, we conducted an ANOVA predicting the proportion of object next mentions given completion type (explanation vs. non-explanation) and verb type (IC vs non-IC). The rate of object re-mention was larger in explanation completions than non-explanation completions (main effect of completion type, significant by subjects, marginal by items: $F_1(1,37)=5.636, p<0.05; F_2(1,18)=4.130, p=0.06$). The rate of object re-mentions was larger following IC verbs than non-IC verbs (main effect of verb type: $F_1(1,37)=42.187, p<0.001; F_2(1,18)=42.736, p<0.001$). There was also a completion type × verb type interaction, whereby object re-mentions were most frequent in explanation completions following IC verbs (interaction: $F_1(1,27)=92.67, p<0.001; F_2(1,15)=42.947, p<0.001$). Because the data for the four conditions is not balanced (e.g., there were only 70 explanations following non-IC verbs but 219 explanations following IC verbs), a mixed-effect logistic regression is useful for evaluating main effects and the interaction between completion type and verb type. In the regression, verb type and completion type were significant factors for modeling the binary outcome of choice of next mention, and the completion type × verb type interaction was also significant ($\beta_{verbType}=1.239, p<0.001; \beta_{completionType}=0.782, p<0.001; \beta_{interaction}=3.518, p<0.001$). Pairwise comparisons show that the rate of object re-mention in explanations following IC verbs (67.2%) is larger than that in explanations following non-IC verbs (8.6%; $t(69)=10.327, p<0.001$); it is larger than that in non-explanations following IC verbs (20.9%; $t(70)=7.728, p<0.001$), and it is larger than that in all other contexts collapsed together (20.2%; $t(82)=6.878, p<0.001$). The appendix lists the individual coherence and next-mention biases for each item, showing that the IC biases are not categorical but rather lie along a continuum (see Caramazza et al., 1977).

### 3.3.3. Results: modeling RTs with empirical item norms

Whereas the analyses reported in Section 3.2 used a dichotomous verbtype predictor (IC or non-IC) to model reading times, here we compare these analyses with new analyses using our norming study results as continuous predictors. We report analyses of our central results—RTs at the first and second spillover regions—using linear mixed-effects models with random subject- and item-specific intercepts. We consider models with the dichotomous verbtype predictor, models with a gradient norm (we investigated both the verb-specific bias toward an upcoming explanation and the bias toward object-NP re-mention given an explanation completion) as a predictor, and models with both as predictors. In all cases, RC attachment and its interactions were included as predictors, and all predictors were centered.

Our analysis using bias towards an upcoming explanation did indeed recover a significant interaction between item-specific bias and RC attachment on
reading times in both spillover regions (first spillover region, explanation bias: $\beta = -2.41, p_{\text{MCMC}} = 0.72$; attachment: $\beta = 6.03, p_{\text{MCMC}} = 0.30$; verb-specific explanation bias $\times$ attachment: $\beta = 14.23, p_{\text{MCMC}} < 0.025$; second spillover region, explanation bias: $\beta = -1.13, p_{\text{MCMC}} = 0.81$; attachment: $\beta = -1.63, p_{\text{MCMC}} = 0.77$; verb-specific explanation bias $\times$ attachment: $\beta = 14.19, p_{\text{MCMC}} < 0.025$). We investigated whether our gradient-bias measure gives us additional explanatory power above and beyond our categorical verb-type predictor—and vice-versa—by fitting a model with both predictors and using likelihood-ratio tests to compare it with models with only the categorical verb-type predictor and only the gradient-bias predictor (in all cases, RC attachment and its appropriate interactions with the bias predictors were included). In neither case did the likelihood-ratio test indicate a significant improvement of the model with both predictor types over either model with only one predictor type (log-likelihood in the first spillover region: $-7278.7$ with both predictors, $-7278.9$ with only categorical verbtype; $-7279.8$ with only quantified bias; in the second spillover region, $-7170.06$ with both predictors, $-7171.1$ with only categorical verbtype, $-7170.11$ with only quantified bias). This result presumably reflects the fact that the two predictors are strongly correlated ($r = 0.773, t(1094) = 40.327, p < 0.001$). The fact that quantified bias toward upcoming explanation does not improve significantly on categorical verbtype as a predictor of reading times could easily be due to the measurement error inherent in estimating explanation bias from a continuation-study sample of limited size.

When using bias toward object-NP re-mention given an explanation, the results are nearly identical in the first spillover region, where we recover a significant interaction between quantified bias and RC attachment, but less clear in the second spillover region, where the numerical direction of the interaction is correct, but fails to reach significance (first spillover region, next-mention bias: $\beta = -8.76, p_{\text{MCMC}} = 0.22$; attachment: $\beta = 4.93, p_{\text{MCMC}} = 0.47$; next-mention bias $\times$ attachment: $\beta = 17.81, p_{\text{MCMC}} < 0.05$; second spillover region, next-mention bias: $\beta = -8.11, p_{\text{MCMC}} = 0.18$; attachment: $\beta = -1.24, p_{\text{MCMC}} = 0.83$; next-mention bias $\times$ attachment: $\beta = 6.45, p_{\text{MCMC}} < 0.27$). Note that four non-IC verbs yielded no explanation continuations so the object-NP re-mention bias for explanations could not be calculated; data for those items was not included. Likelihood-ratio tests indicate that in the first and second spillover regions, the combined model is statistically indistinguishable from either the quantified-bias or categorical-verbtype model (log-likelihood in the first spillover region: $-5811.7$ with both predictors, $-5812.9$ with only categorical verbtype; $-5811.9$ with only quantified bias; in the second spillover region, $-5703.7$ with both predictors, $-5704.3$ with only categorical verbtype, $-5704.3$ with only quantified bias). As with explanation bias, re-mention bias is strongly correlated with verbtype ($r = 0.858, t(875) = 49.448, p < 0.001$); we attribute the poorer explanatory power
of re-mention bias to the smaller sample size used to compute it (see Table 2), and the concomitantly larger measurement error.

3.4. Discussion

Experiment 2 was designed to test the hypothesis that expectations about discourse continuations have an impact on the immediate processing of RCs. This hypothesis was confirmed by the significant interaction between attachment level and verbtype on reading times at the first and second spillover regions immediately after the disambiguating finite verb in the relative clause. In fact, not only was the bias toward low-attaching RCs reduced in the IC condition as compared to the non-IC condition (as predicted), but analyses within the IC condition at the first spillover region and across the first two spillover regions showed a reversal of the default low-attachment bias, such that high-attaching RCs were actually read more quickly than low-attaching RCs. Because the operative biases are tied only to comprehenders’ expectations for an ensuing explanation rather than the actual existence of one – after all, at the spillover region the information needed to establish that an explanation relation indeed exists has yet to be encountered – these results stand in stark opposition to clausal-integration accounts. Despite the fact that several different discourse-level sources of information contribute – expectations about ensuing coherence relations, biases toward next mention conditioned on coherence relations, and knowledge that RCs can provide explanations – and further that their integration requires a fairly complex chain of reasoning, the timecourse of the effect suggests that participants are invoking these biases mid-sentence and using them to generate expectations about upcoming syntactic structure.

4. General Discussion

As we described in the introduction, Hume argued that there are fundamental principles of association that serve to connect ideas in the human mind. Since it is human minds that produce and interpret natural language, we would expect, and do in fact find, that utterances in coherent discourses should likewise be connected by such associative principles. It has also been argued that these principles in turn affect the manner in which discourse-dependent linguistic forms (e.g., ellipsis, anaphora, intonation, tense) are interpreted (Hobbs, 1979; Venditti et al., 2002; Kehler, 2005; Asher and Lascarides, 2003; Kaiser, 2009, inter alia). Recent work has further shown that comprehenders not only infer such coherence relations, but also generate expectations concerning what relations are likely to ensue and that these expectations affect how pronominal forms are processed (Rohde et al., 2006, 2007; Kehler et al., 2008).
The work reported here took this idea further by asking whether discourse coherence expectations influence local syntactic processing decisions within a sentence. Because syntactic processing is generally considered to be driven by interpretation mechanisms that differ from those that drive discourse comprehension, there has been no obvious reason to expect that syntactic processing would be sensitive to discourse coherence expectations. However, by virtue of an integrative design that examined comprehenders’ behavior when processing sentences involving RC attachment ambiguity, the results reported herein demonstrate that syntactic processing is indeed sensitive to these types of expectations. Our sentence-completion study showed first that object-biased IC verbs yielded RCs that are more likely to explain the event in the matrix clause than RCs following non-IC verbs, and second, that these explanation-providing RCs are more likely to attach high. These results are in keeping with previous work on intersentential coherence (Kehler et al., 2008) suggesting that the presence of an IC verb influences comprehenders’ expectations about the direction the discourse is likely to take – raising the expectation for an upcoming explanation and creating a concomitant bias for a subsequent mention of the causally-implicated referent. Our self-paced reading study yielded a reversal of the typically-observed RC low-attachment preference in cases in which a referent that is causally-implicated by an IC verb was located at the high attachment site. These on-line results suggest that comprehenders track such expectations mid-sentence, and that these expectations are used incrementally to influence syntactic processing. This novel and surprising result is explainable only by assuming that comprehenders are aware that the class of IC verbs used in our experiments impute causality to their direct object, that they know that clauses with IC verbs are likely to be followed by an explanation of the eventuality the clause denotes, that they are aware that RCs can be used to implicate explanations, and, crucially, can bring all of these pieces of information together so as to influence an on-line syntactic attachment decision.

Our use of a design centered around RC processing is fitting given that the field of psycholinguistics has already placed considerable emphasis on RCs as a testing ground for a variety of phenomena, including ambiguity resolution (Cuetos and Mitchell, 1988), extraction (Kluender, 1992), optionality (Ferreira and Dell, 2000), expectation-based parsing (Jurafsky, 1996; Hale, 2001; Levy, 2008), and thematic fit (McRae et al., 1998), among others. In the remainder of this section, we situate our results relative to previous work on RC attachment ambiguity and discourse-sensitive models of sentence processing.
4.1. RC processing across languages and discourse contexts

Our study maintains continuity with a line of research begun by Cuetos and Mitchell (1988) showing that the Late Closure strategy of low postmodifier attachment, once hypothesized to be a principled and cross-linguistically universal syntax-first parsing preference (Frazier, 1987), is not in fact consistent across languages or even across contexts within a language (see reviews in Cuetos et al., 1996; Mitchell and Brysbaert, 1998; Desmet et al., 2002; and Papadopoulou and Clahsen, 2006). Studies involving ambiguously attached RC constructions such as that in (3) (repeated here as (23)) in languages including Spanish, French, German, and Dutch (among others) have found evidence of a bias towards a high-attachment interpretation (Cuetos and Mitchell, 1988; Zagar et al., 1997; Hemforth et al., 2000; Brysbaert and Mitchell, 1996).

(23) Someone shot the servant of the actress who was on the balcony.

The lack of a universal attachment preference has thus been problematic for theories of sentence processing which posit crosslinguistic syntactic constraints and strategies. Moreover, it has also been shown that within a single language, attachment preferences vary with lexical properties, such as animacy (Desmet et al., 2006), referentiality (Gilboy et al., 1995), and even specific lexical head (Desmet and Gibson, 2003). Evidential accounts such as the competition-integration model (MacDonald, 1994; Spivey and Tanenhaus, 1998; McRae et al., 1998) and probabilistic models (Jurafsky, 1996; Narayanan and Jurafsky, 1998, 2002; Crocker and Brants, 2000; Hale, 2001; Levy, 2008) are compatible with the idea that a host of factors involving multiple information sources conspire to determine attachment preferences in any particular sentence. In this view, so-called “default” attachment preferences in a given language are simply the consequence of the distribution of relevant information-source particulars in the language.

There has also been some work on the role of discourse processing in on-line sentence comprehension. Perhaps the earliest account relating discourse-level processing and syntactic disambiguation was the referential theory (Crain and Steedman, 1985; Altmann and Steedman, 1988), which focused on the ability of NP postmodifiers to restrict the domain of possible reference of the modified NP. According to this theory, an NP with a restrictive postmodifier such as the horse raced past the barn can, in a typical discourse context, be taken not only to presuppose the existence of a horse that was raced past a barn, but also to conversationally implicate (Grice, 1975) the existence of a horse that was not. This implicature results, according to Gricean reasoning, from that fact that if there were only one horse in the context, the speaker would be expected to have chosen the less informative and less prolix NP the horse. As a result, when there
is ambiguity as to whether material after a given NP constitutes a postmodifier of that NP, the postmodifier analysis should be favored when the preceding context implies that the NP would otherwise be referentially ambiguous. It has been argued by some researchers that it is implausible to expect that the inference of a conversational implicature could affect on-line syntactic comprehension, e.g.:

To make a conversational implicature, a listener must have already parsed the sentence, assigned it its literal interpretation, realised that additional inferences must be added to make it conform to the Gricean maxim, and determined what these inferences are. Such activity could not reasonably affect the initial steps of parsing. (Clifton and Ferreira, 1989)

Nevertheless, it has since been shown by Ni et al. (1996) and Sedivy (2002) that invoking implicit referential contrast sets can affect the main-verb/reduced-relative ambiguity in classic garden-path sentences. More generally, one might predict that the manipulation of the number of compatible referents available would affect the degree to which one would expect a post-modifier to appear with definite referential NPs. This prediction was verified by Van Berkum et al. (1999), who showed in a Dutch-language ERP experiment that comprehenders reading sentence onsets of the type David told the girl that . . . , where that is ambiguous between introducing a complement clause versus a relative clause, had stronger relative-clause expectations in an ambiguous (two-girl) referential context than in an unambiguous (one-girl) referential context. Carrying over the same logic to RC attachment ambiguity, one might also predict that RC attachment to a definite NP would be preferred when the NP would be referentially ambiguous (and hence infelicitous) without a postmodifier. In (23), for example, one might predict a low-attachment bias in a context with multiple actresses as in (24) below, but a high-attachment bias in a context with multiple servants as in (25):

(24) There was a servant working for two actresses. Someone shot the servant of the actress who was on the balcony.

(25) There were two servants working for a famous actress. Someone shot the servant of the actress who was on the balcony.

There has been evidence supporting this prediction from research using off-line methodologies on a variety of languages (French: Zagar et al., 1997; Dutch: Desmet et al., 2002; Greek: Papadopoulou and Clahsen, 2006). The evidence for on-line effects has been more mixed. Papadopoulou and Clahsen (2006) reported significant effects using self-paced reading, but the Zagar et al. (1997) and Desmet
et al. (2002) studies found no significant effects using eye-tracking.

Our experiments complement these existing studies, but go beyond them in two key respects. First, ours is the first study to look at the effect of discourse processing beyond referential ambiguity on syntactic comprehension. Second, in response to question #3 posed in the introduction, the modulation of attachment preferences observed in our study cannot be reduced to a simple bias against infelicity in definite descriptions (or in any other aspect of the sentence). For example, (15a), repeated below as (26), is in no way infelicitous despite the fact that the RC does not provide an explanation of the detesting; the need for an explanation can easily be satisfied by a subsequent sentence (27).

(26) John detests the children of the musician who lives in La Jolla.

(27) The children are arrogant and rude.

That is, even though an explanation RC continuation is not required by the context set up by the words

(28) John detests the children of the musician who...

for utterance felicity, this context nevertheless induces a probabilistic expectation that the RC will in fact be an explanation, which in turn induces a high RC attachment preference associated with effects that can be measured online. Taken together, these results suggest that coherence-driven factors are as important as, and fully integrated with, lexical and morphosyntactic cues in resolving ambiguity. Unlike previous results which are consistent with models in which discourse context might only come into play to ward off impending infelicity or ungrammaticality, the coherence-driven discourse biases we present here are active in general contexts in which felicity and grammaticality are not at stake, i.e., in most common discourses.

We have taken advantage of the fact that an RC can be used to express an explanation for a matrix-clause event in modulating attachment preferences. Notably, as described in Section 1.4, such explanations are only pragmatically implicated when provided by an RC – and as such, our study demonstrates that a comprehender’s expectations about ensuing conversationally-implicated information can indeed influence the initial steps of parsing (cf. the foregoing quote from Cliffton and Ferreira, 1989). The experiments also establish a potentially intriguing connection between the next mention biases in Explanation relations that Kehler et al. (2008) found for pronoun interpretation and those illustrated here for RC attachment, which itself can be seen as involving the anaphoric bind-
ing of a *wh*-pronoun (see Hemforth et al., 2000 for discussion of the relationship between pronoun interpretation and *wh*-pronoun binding). We now turn to the question of what types of on-line processing models can account for the results reported in this paper.

4.2. Expectation-Based Models for Discourse-Sensitive Syntactic Comprehension

There are several aspects of our experimental results that any processing model needs to account for. First, there was a close match between off-line attachment preferences in sentence completion in Experiment 1 and on-line reading preferences in Experiment 2: In both cases, the presence of an object-biased IC verb eliminated low-attachment bias. Experiment 1 makes it clear that this shift in attachment preference is driven by a much higher proportion of RC completions providing an explanation for the matrix-clause event in the IC condition than in the non-IC condition. Second, the on-line preferences in Experiment 2 emerged at the first spillover region after the disambiguating RC verb (i.e., *is/are* in (20)–(21)), well before participants encountered the lexical material necessary to establish that the RC provides an explanation for the matrix-clause event. These results are problematic for any theory in which the establishment of inter-clausal discourse coherence relations (including those between matrix and relative clauses) is not fully incremental.

The question remains as to what theories can naturally capture these results, and how. We believe that these results can be handled by a range of evidential models of on-line comprehension, so long as (a) syntactic attachment preferences are made fully incrementally (contra the delay model of Rayner and Frazier (1987)), and (b) discourse-based preferences are taken into account and can interact fully with any other biases that may be active. These models include at least the competition-integration model (MacDonald, 1994; Spivey and Tanenhaus, 1998; McRae et al., 1998), probabilistic disambiguation/pruning and attention-shift models (Jurafsky, 1996; Narayanan and Jurafsky, 1998, 2002; Crocker and Brants, 2000), and surprisal (Hale, 2001; Levy, 2008). Although none of these models as presented to date have explicitly included discourse constraints, their probabilistic architectures allow for the incorporation of potentially arbitrary information sources, and thus there is no reason why discourse factors could not be smoothly incorporated into any of them. As an illustration, Appendix B provides a formal probabilistic model of how knowledge about coherence relations could be deployed to modulate incremental RC attachment preferences. In this model, the crucial quantity determining comprehension difficulty at the disambiguating RC verb is the probability of high versus low attachment given the preceding context. It is assumed that comprehenders have access to probability distributions over (i) inter-clausal coherence relations given the first clause (the
main clause in our experiments), (ii) next-mention biases given first clause and the inter-clausal coherence relation with the next clause, and (iii) the probability of expressing the next clause as an RC (as opposed, for example, to expression as a new sentence). In determining the online RC attachment biases in our experiment, this model consults distributions (ii) and (iii), marginalizing over the coherence-relation distribution (i).

4.3. The Immediate Focusing vs. Clausal Integration Debate

Our results also weigh in on the recent controversy concerning the time course during which IC information is used in sentence processing, a debate that has until now been centered on its use in pronoun interpretation. Proponents of the immediate focusing account (McKoon et al., 1993; McDonald and MacWhinney, 1995; Koornneef and Van Berkum, 2006, inter alia) argue that IC biases are utilized early enough so as to essentially constitute a focusing mechanism when interpreting pronominal expressions. Proponents of the clausal integration account (Garnham et al., 1996; Stewart et al., 2000, inter alia), on the other hand, argue that IC information is used only as part of a sentence-final clause integration process. The clausal integration account predicts that IC effects will arise later during sentence interpretation than the immediate focusing account does, at least when a pronoun occurs early in the clause.

Our results strongly support the immediate focusing account. As we have argued, our results suggest that comprehenders (i) know that the class of IC verbs used in our experiments impute causality to their direct object, (ii) know that IC verbs create a strong expectation that an explanation of the event described by a matrix IC clause is soon to follow, (iii) know that RCs can be used to implicate explanations, and, crucially, (iv) bring all of these pieces of information together so as to influence an on-line syntactic attachment decision. If IC information – not only IC biases toward a particular referent, but also the biases they contribute toward expecting an explanation to ensue – is not utilized until downstream linguistic material in the RC is processed, we are left with no explanation for why we find an effect in the spillover region immediately after the disambiguating auxiliary.

Importantly, however, we are not suggesting that information occurring later in the RC will have no impact on syntactic attachment effects. In an analysis such as ours in which expectations are updated on an incremental, word-by-word basis, information encountered at any time can force a revision to the probabilities assigned to particular syntactic decisions. For instance, subsequent words in the RC might reduce the likelihood that the RC expresses an explanation, which would in turn reduce the likelihood of a high attachment. Alternatively, subsequent
words might reinforce the expectation of an explanation, but one that mentions the low-position NP instead of the high one. In such situations, the model would predict processing difficulty at the time (or soon after) the bias-incongruent information is encountered. The crucial point, however, is that comprehenders do not wait until downstream in the clause to start utilizing IC-driven probabilistic expectations; they instead use expectations based on information available at the time the RC is encountered (IC and otherwise) in making syntactic attachment decisions.

This view is supported by recent IC studies of pronoun interpretation. To take one example, Koornneef and Van Berkum (2006) looked for mid-sentence reading delays caused by pronouns that are inconsistent with the bias of a preceding IC verb in two experiments with gender-unambiguous pronouns. In a word-by-word self-paced reading task, they found that words in the pre-critical region were read equally fast across the bias-consistency conditions, but readers slowed down right at a bias-inconsistent pronoun, with a significant main effect emerging at the first two words thereafter. In an eye tracking study that measured mean regression path durations, again no differences were measured in the pre-critical region, but pronouns that were inconsistent with the IC bias reliably perturbed the reading process at or shortly after the pronoun. The results of both experiments therefore suggest that IC information becomes available rapidly enough to appear mid-sentence, even in passages in which the gender of the pronoun singles out a unique referent. Similarly, Pyykönen and Järvikivi (in press) ran an eye-tracking study in Finnish using a visual-world paradigm that asked whether IC effects would occur before the end of a clause containing an IC verb, that is, before either a connective or pronoun was encountered. They found a significant main effect of IC starting at 900 ms after the verb onset, i.e., just after participants had encountered the verb/subject/object complex. These studies therefore support immediate focusing accounts. The current study does as well, and in fact does so by extending the range of IC-sensitive phenomena to include relative clause attachment in addition to pronoun interpretation.

4.4. Discourse Continuations as a Unit of Prediction

Although our results can be incorporated into immediate-focusing incremental models of probabilistic syntactic comprehension, they further constrain these models in terms of the information sources that they must include. That is, models of sentence processing can no longer be built separately from models of discourse processing. Whereas most researchers would not deny that complex inferencing affects language usage and comprehension, finding a quantifiable unit over which to estimate predictions for processing effects has been hard to do. In this work, we have offered coherence relations as a concrete, quantifiable feature
of the discourse context with which probabilistic expectations can be calculated.\footnote{In the analysis in the Appendix, for example, intraclausal discourse coherence relations are both a unit of prediction — \(P(\text{Reln}|M)\) — and a feature implicated indirectly in a syntactic prediction — \(P(\text{NP}_1|\text{Reln}, M)\) and \(P(\text{RC high}|\text{NP}_1, \text{Reln}, M)\).} An advantage of the approach taken in this paper is that a concrete factor (verb-type) was manipulated in order to generate repercussions at the level of coherence. As such, ‘discourse-level factors’ need not be relegated to the status of haphazard or fuzzy cues (see Kadmon, 2001 for a discussion of what constitutes a pragmatic explanation) nor do linguistic studies need to restrict the analyses to ‘neutral’ discourse contexts and make claims concerning processing biases that are active \textit{all else being equal}. In fact, all else is never equal, and our hope is that acknowledging this lack of neutrality in the discourse context will lead to more research quantifying the properties and structure of the surrounding discourse.

5. Conclusions

Fundamental principles of association – characterized here by coherence relations – serve as the glue that connect ideas expressed in coherent discourses. The studies reported in this paper have established that certain contexts increase the likelihood of certain discourse coherence relations and that these biases have an impact on RC attachment ambiguity. Expectations regarding upcoming coherence relations were shown to arise from properties of the matrix clause, specifically the matrix verb. By varying the type of verb in the matrix clause, we found that a subsequent RC completion was more likely to provide an explanation of the event in the matrix clause if the verb belonged to the class of IC verbs (Experiment 1). The RC was also more likely to attach to the direct object of an object-biased IC verb than to the object of a non-IC verb in cases in which the direct object occupied the higher position of a complex NP. We also found that these off-line preferences were mirrored in on-line comprehension, as evidenced by the interaction between RC attachment level and matrix-clause verb-type in our self-paced reading study—even before the bottom-up input indicating the true coherence relation between matrix and relative clauses was available (Experiment 2). Since these effects occur before comprehenders have been exposed to complete clauses, we conclude that expectations about interclausal discourse coherence relations are updated fully incrementally, and have moment-by-moment influence on syntactic disambiguation. Existing models that have been proposed to account for the widely reported dispreference for high attachments in English fail to account for these results since, in our experiments, the structure of the complex NP, the lexical properties of the individual nouns, and the referential context were the same across conditions. What changed between conditions
was the expectation triggered by the matrix clause verb regarding the likelihood of an upcoming Explanation relation. The fact that comprehenders appear to be sensitive to coherence-level biases mid-sentence attests to the importance of constructing models of sentence processing that incorporate information about discourse coherence relations.

Acknowledgments

This research was supported by NIH grant T32-DC000041 to the Center for Research in Language at the University of California, San Diego and NSF IGERT grant DGE-0333451. The results from this paper have been presented at the 21st CUNY Conference on Human Sentence Processing, Chapel Hill, NC, March 2008. We also thank Klinton Bicknell, Laura Kertz, Meredith Larson, and research assistant Erica Gold.

References


38


Jaeger, T. F. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language (Special issue on Emerging Data Analysis) 59:434–446.


cations.
Dordrecht: Kluwer.
Pyykkönen, P., and J. Järviäivi. in press. Activation and persistence of implicit causality information in spoken language comprehension. Experimental Psychology.


A. Experimental materials

Story completion stimuli (Experiment 1)
The stimuli were all of the form Name - IC/non-IC verb - complex NP - who.

1. Carl admires/works with the agent of the rockstars who...
2. Greg adores/smiles at the secretaries of the lawyer who...
3. Jared blamed/noticed the friends of the athlete who...
4. Frank complimented/met the guests of the bride who...
5. Bill congratulated/visited the teacher of the second-graders who...
6. Candice criticized/talked to the leader of the activists who...
7. Beth despisest/babysits the children of the jazz musician who...
8. Casey detests/looks like the father of the students who...
9. Melissa dislikes/watches the little girls of the neighbor who...
10. Sandra insulted/chatted with the gardeners of the millionaire who...
11. Ryan likes/resembles the captain of the old sailors who...
12. Joel pities/hires the bodyguards of the celebrity who...
13. Ken praised/videotaped the assistants of the CEO who...
14. Alan punished/saw the accountant of the businessmen who...
15. Tina resents/knows the doctors of the supermodel who...
16. Luis recognized/scolded the landlady of the actors who...
17. Craig rewarded/inspected the servants of the dictator who...
18. Scott ridiculed/counted the fans of the singer who...
19. George thanked/interviewed the representative of the employees who...
20. Alice values/lives next to the surgeon of the soldiers who...
21. Paul worships/listens to the coach of the cheerleaders who...
Sample story completions (taken from participants’ completions from Experiment 1)

1. Beth babysits the children of the jazz musician who lives in La Jolla.  
   [NON-IC, NON-EXP, LOW]
2. Frank met the guests of the bride who were her friends from high school.  
   [NON-IC, NON-EXP, HIGH]
3. Melissa watches the little girls of the neighbor who works evening shifts.  
   [NON-IC, EXP, LOW]
4. Craig inspected the servants of the dictator who were suspected of stealing.  
   [NON-IC, EXP, HIGH]
5. Melissa dislikes the little girls of the neighbor who lives on her right.  
   [IC, NON-EXP, LOW]
6. Frank complimented the guests of the bride who were sitting in the front row.  
   [IC, NON-EXP, HIGH]
7. Bill congratulated the teacher of the second-graders who had all learned their times tables.  
   [IC, EXP, LOW]
8. Alan punished the accountant of the businessmen who was notorious for IRS fraud.  
   [IC, EXP, HIGH]
Verb biases observed in Experiment 1

Verbs differed in the proportion of explanation RCs produced.

<table>
<thead>
<tr>
<th>Verb</th>
<th>Class</th>
<th>% Expl</th>
<th>Verb</th>
<th>Class</th>
<th>% Expl</th>
</tr>
</thead>
<tbody>
<tr>
<td>chat-with</td>
<td>Non-IC</td>
<td>0%</td>
<td>hire</td>
<td>Non-IC</td>
<td>34.8%</td>
</tr>
<tr>
<td>count</td>
<td>Non-IC</td>
<td>0%</td>
<td>adore</td>
<td>IC</td>
<td>36.0%</td>
</tr>
<tr>
<td>interview</td>
<td>Non-IC</td>
<td>0%</td>
<td>insult</td>
<td>IC</td>
<td>36.4%</td>
</tr>
<tr>
<td>know</td>
<td>Non-IC</td>
<td>0%</td>
<td>watch</td>
<td>Non-IC</td>
<td>46.2%</td>
</tr>
<tr>
<td>live-next-to</td>
<td>Non-IC</td>
<td>0%</td>
<td>compliment</td>
<td>IC</td>
<td>50.0%</td>
</tr>
<tr>
<td>look-like</td>
<td>Non-IC</td>
<td>0%</td>
<td>praise</td>
<td>IC</td>
<td>50.0%</td>
</tr>
<tr>
<td>meet</td>
<td>Non-IC</td>
<td>0%</td>
<td>admire</td>
<td>IC</td>
<td>52.6%</td>
</tr>
<tr>
<td>recognize</td>
<td>Non-IC</td>
<td>0%</td>
<td>reward</td>
<td>IC</td>
<td>54.5%</td>
</tr>
<tr>
<td>resemble</td>
<td>Non-IC</td>
<td>0%</td>
<td>scold</td>
<td>IC</td>
<td>60.0%</td>
</tr>
<tr>
<td>see</td>
<td>Non-IC</td>
<td>0%</td>
<td>videotape</td>
<td>Non-IC</td>
<td>61.1%</td>
</tr>
<tr>
<td>talk-to</td>
<td>Non-IC</td>
<td>0%</td>
<td>blame</td>
<td>IC</td>
<td>64.3%</td>
</tr>
<tr>
<td>work-with</td>
<td>Non-IC</td>
<td>0%</td>
<td>criticize</td>
<td>IC</td>
<td>66.7%</td>
</tr>
<tr>
<td>visit</td>
<td>Non-IC</td>
<td>5.3%</td>
<td>dislike</td>
<td>IC</td>
<td>76.2%</td>
</tr>
<tr>
<td>babysit</td>
<td>Non-IC</td>
<td>8.0%</td>
<td>worship</td>
<td>IC</td>
<td>80.0%</td>
</tr>
<tr>
<td>smile-at</td>
<td>Non-IC</td>
<td>8.7%</td>
<td>congratulate</td>
<td>IC</td>
<td>81.0%</td>
</tr>
<tr>
<td>inspect</td>
<td>Non-IC</td>
<td>13.0%</td>
<td>despise</td>
<td>IC</td>
<td>82.6%</td>
</tr>
<tr>
<td>notice</td>
<td>Non-IC</td>
<td>14.3%</td>
<td>pity</td>
<td>IC</td>
<td>82.6%</td>
</tr>
<tr>
<td>listen-to</td>
<td>Non-IC</td>
<td>18.2%</td>
<td>resent</td>
<td>IC</td>
<td>84.0%</td>
</tr>
<tr>
<td>value</td>
<td>IC</td>
<td>22.7%</td>
<td>punish</td>
<td>IC</td>
<td>85.7%</td>
</tr>
<tr>
<td>like</td>
<td>IC</td>
<td>26.7%</td>
<td>ridicule</td>
<td>IC</td>
<td>91.3%</td>
</tr>
<tr>
<td>thank</td>
<td>IC</td>
<td>33.3%</td>
<td>detest</td>
<td>IC</td>
<td>95.7%</td>
</tr>
</tbody>
</table>

Reading time stimuli (Experiment 2)

The stimuli were all of the form Name - IC/non-IC verb - complex NP - who - singular/plural RC verb - adverb - continuation. The forward slash (‘/’) separates alternatives that differed between conditions (IC/non-IC; singular/plural). An underscore (‘_’) connects words that were revealed together in one region. Comprehension questions are listed in brackets.

1. Anna scolded/studied with the chef of the aristocrats who was/were routinely letting food go to waste. [Did food go to waste?]
2. John stared at/lived next to the teacher of the second graders who was/were definitely smartest in the school. [Was the teacher/the second graders smart?]
3. Jenny assisted/joked with the maid of the executives who was/were regularly late to work. [Were the executives/was the maid late to work?]
4. Nick trusted/stood near the captain of the sailors who has/have consistently weathered big storms. [Did the captain have Nick’s confidence? Was Nick near the captain?]

5. Angela corrected/gossiped with the secretary of the lawyers who has/have occasionally made small mistakes. [Have there been occasional errors?]

6. Bob comforted/greeted the leader of the activists who was/were deeply disappointed by the court’s decision. [Was Bob disappointed with the court’s decision?]

7. Laura envies/knows the manager of the cashiers who has/have supposedly received a huge raise. [Did the manager/cashiers get a huge raise?]

8. Zack valued/recognized the daughter of the shopkeepers who was/were usually willing to spot him a few dollars. [Did Zack lend money to the daughter?]

9. Sarah fears/jogs with the uncle of the toddlers who is/are often heard yelling and screaming. [Are toddlers known for being well behaved?]

10. Adam noticed/resembled the representative of the employees who was/were always wearing safety goggles. [Were the employees / Was the representative wearing safety goggles?]

11. Tina praised/met the gardeners of the millionaire who has/have recently installed a solar powered sprinkler. [Has the millionaire / Have the gardeners put in a new sprinkler system?]

12. Justin hates/carpools with the cousins of the accountant who is/are forever telling the same tasteless jokes. [Is the accountant / Are the cousins likeable?]

13. Emily blamed/waited with the nieces of the florist who has/have repeatedly ruined expensive orchids. [Did some flowers get damaged?]

14. Joe helped/ran into the brothers of the athlete who is/are perpetually failing math class. [Are the brothers / Is the athlete failing math?]

15. Jessica reproached/worked with the doctors of the supermodel who was/were adamantly in favor of plastic surgery. [Did the supermodel/doctors advocate plastic surgery?]

16. Brian pacified/visited the associates of the businessman who was/were nearly bankrupted by the new tax policy. [Did the new tax policy benefit businesses?]

17. Melissa detests/babysits the children of the musician who is/are generally arrogant and rude. [Does Melissa get frustrated with the children? / Could Melissa be a teenager?]

18. Frank thanked/talked to the servants of the dictator who has/have lately been helping the poor. [Does Frank admire altruism? / Did Frank talk to the dictator’s staff?]
19. Tracy congratulated/chatted with the bodyguards of the celebrity who was/were constantly fighting off the paparazzi. [Does the paparazzi ignore celebrities?]
20. Kevin mocked/counted the fans of the singer who was/were continually stagediving and getting hurt. [Were the fans diving off the stage? Is the singer someone who dives off the stage?]

**Item biases observed in Norming study for Experiment 2 (by verb *v*)**

| Item                                                    | *p(expl|v)* | *p(object|v,expl)* | *p(object|v,non-expl)* |
|---------------------------------------------------------|----------|-----------------|----------------------|
| scold/chef/aristocrats                                  | 0.93     | 0.64            | 0.00                 |
| study-with/chef/aristocrats                             | 0.50     | 0.22            | 0.11                 |
| stare/teacher/2nd-graders                               | 0.70     | 0.21            | 0.33                 |
| live-next-door-to/teacher/2nd-graders                   | 0.00     | n.a.            | 0.14                 |
| assist/maid/executives                                  | 0.62     | 0.25            | 0.40                 |
| joke-with/maid/executives                               | 0.14     | 0.00            | 0.50                 |
| trust/captain/sailors                                   | 0.68     | 0.85            | 0.17                 |
| stand-near/captain/sailors                              | 0.27     | 0.25            | 0.00                 |
| correct/secretary/lawyers                               | 0.53     | 0.63            | 0.43                 |
| gossip-with/secretary/lawyers                           | 0.18     | 0.50            | 0.11                 |
| comfort/leader/activists                                | 0.46     | 0.50            | 0.14                 |
| greet/leader/activists                                  | 0.20     | 0.00            | 0.08                 |
| envy/manager/cashiers                                   | 0.83     | 0.47            | 0.00                 |
| know/manager/cashiers                                   | 0.40     | 0.00            | 0.11                 |
| value/daughter/shopkeeper                               | 0.25     | 1.00            | 0.07                 |
| recognize/daughter/shopkeeper                           | 0.32     | 0.16            | 0.08                 |
| fear/uncle/toddlers                                     | 0.78     | 0.79            | 0.00                 |
| jog-with/uncle/toddlers                                 | 0.19     | 0.00            | 0.31                 |
| notice/representative/employees                          | 0.44     | 1.00            | 0.40                 |
| resemble/representative/employees                        | 0.00     | n.a.            | 0.00                 |
| praise/gardeners/millionaire                             | 0.75     | 0.56            | 0.33                 |
| meet/gardeners/millionaire                               | 0.05     | 0.00            | 0.53                 |
| hate/cousins/accountant                                 | 0.70     | 1.00            | 0.00                 |
| carpool-with/cousins/accountant                          | 0.40     | 0.25            | 0.33                 |
| blame/nieces/florist                                    | 0.79     | 0.91            | 0.00                 |
| wait-with/nieces/florist                                | 0.33     | 0.00            | 0.38                 |
B. Probabilistic analysis of Experiment 2

We assume that the difficulty at the first spillover region in Experiment 2 is indeed spillover generated by attachment-level disambiguation at the preceding RC verb. Hence, we focus on the conditional probability of high versus low RC attachment immediately before this verb. For ease of illustration, we will cast our formulation specifically in terms of examples (20) and (21), repeated below as (29) and (30):

**(29)** NON-IC MATRIX: John babysits the children of the musician ...
   a.  [LOW ATTACHMENT] ... who is generally arrogant and rude.
   b.  [HIGH ATTACHMENT] ... who are generally arrogant and rude.

**(30)** IC MATRIX: John detests the children of the musician ...
   a.  [LOW ATTACHMENT] ... who is generally arrogant and rude.
   b.  [HIGH ATTACHMENT] ... who are generally arrogant and rude.

All simplifications that result from focusing on (29) and (30) are made without loss of generality.

First we introduce some notation for several events:

- **RC {high, low}:** the presence of a subject-extracted relative clause (SRC) that attaches high or low into a two-level NP
• **M**: the context preceding the word *who*, i.e., the matrix clause

• **C**: the context preceding the RC verb, i.e., the word *who* and *M*

• **NP**: the event that the discourse referent mentioned as the subject of the clause following *M* (whether as an SRC or new sentence) will be the *i*-th NP of the NP complex. (In the case that an SRC follows, *NP* is taken to be the referent of *who*.)

• **Reln**: the coherence relation holding between *M* and the next clause

Under this notation, the attachment preference immediately before disambiguation can be expressed by the probability ratio

\[
\frac{P(\text{RC high}|C)}{P(\text{RC low}|C)} \quad (\text{III})
\]

A high-attaching SRC implies the event *NP*—i.e. that the next-mentioned referent is the high NP—and likewise a low-attaching SRC implies *NP*. We also take advantage of the fact that *C* is composed of the events *M* and *who*, giving us

\[
\frac{P(\text{RC high}|C)}{P(\text{RC low}|C)} = \frac{P(\text{RC high}, NP_1|\text{who}, M)}{P(\text{RC low}, NP_2|\text{who}, M)} \quad (\text{IV})
\]

Applying Bayes' Rule, we can rewrite the numerator \(P(\text{RC high}, NP_1|\text{who}, M)\) as \(P(\text{RC high}, NP_1, \text{who}|M)\)/\(P(\text{who}|M)\), and similarly for the denominator, giving us

\[
\frac{P(\text{RC high}|C)}{P(\text{RC low}|C)} = \frac{P(\text{RC high}, NP_1, \text{who}|M)/P(\text{who}|M)}{P(\text{RC low}, NP_2, \text{who}|M)/P(\text{who}|M)} \quad (\text{V})
\]

\[
= \frac{P(\text{RC high}, NP_1, \text{who}|M)}{P(\text{RC low}, NP_2, \text{who}|M)} \quad (\text{VI})
\]

\[
= \frac{P(\text{who}|\text{RC high}, NP_1, M)P(\text{RC high}, NP_1|M)}{P(\text{who}|\text{RC low}, NP_2, M)P(\text{RC low}, NP_2|M)} \quad (\text{VII})
\]

We introduce our first assumption here: That the probability of using *who* to begin an SRC is approximately independent of the attachment level. This allows us to cancel the first terms of the numerator and denominator, giving us

48
Our second assumption is that the probability of expressing the next clause as an SRC is also approximately independent of whether the high or low NP is the subject of the next clause. This allows us to cancel the first terms once more, and making the marginalization over discourse coherence relations explicit gives us:

\[
\frac{P(\text{RC high}|C)}{P(\text{RC low}|C)} \approx \frac{P(\text{RC high}, NP_1|M)}{P(\text{RC low}, NP_2|M)} \quad \text{(VIII)}
\]

\[
= \frac{P(\text{RC high}|NP_1, M)P(NP_1|M)}{P(\text{RC low}|NP_2, M)P(NP_2|M)} \quad \text{(IX)}
\]

The role of discourse coherence relations now becomes clear. In Experiments 1 and 2, we contrasted conditions with and without object-biased IC verbs in \(M\). It has been independently verified that the presence of such an IC verb causes two things to happen: First, it makes the coherence relation \(Reln\) with the next clause more likely to be an Explanation (Kehler et al., 2008), and second, it makes the next-mentioned NP more likely to be the direct object of that verb when \(Reln\) is an Explanation (Garvey and Caramazza, 1974; Brown and Fish, 1983; Au, 1986; McKoon et al., 1993; Kehler et al., 2008). The first effect corresponds to a rise in \(P(Reln=\text{Explanation}|M)\), and the second effect to a rise in \(P(NP_1|Reln=\text{Explanation}, M)\) and a corresponding drop in \(P(NP_2|Reln=\text{Explanation}, M)\). Both these effects are seen not only in the offline results of Experiment 1 but also in the offline norming study for Experiment 2. Thus, the most important term in the sums of Equation X will be that for which \(Reln=\text{Explanation}\), and this term will be large in the numerator (because \(P(NP_1|Reln=\text{Explanation}, M)\) is large) and small in the denominator (because \(P(NP_2|Reln=\text{Explanation}, M)\) is small). As a result, we can expect the probability ratio \(\frac{P(\text{RC high}|C)}{P(\text{RC low}|C)}\) to be substantially larger in the IC condition than in the non-IC condition. This matches the empirical results of Experiment 2, where we found an interaction between RC attachment and verbtype immediately after the RC verb.

Modeling attachment preferences before seeing the RC verb in this way draws a direct connection between RC processing and earlier literature on the effects of IC verbs. By making two assumptions of approximate probabilistic independence to draw this connection – involving the probabilities of using the word \textit{who}
to mark the RC onset, and expressing the next clause as an RC – our formulation makes totally explicit all the conditions that are required to predict the probabilistically-based processing differences that occurred in our experiment. Importantly, however, we do not claim that these independence assumptions are imposed absolutely by comprehenders, but rather that they are approximately correct for the materials we used. Furthermore, if we constructed materials that broke these independence assumptions, we would expect different patterns of processing difficulty. As an example, if one of the NPs were inanimate, then the second assumption (independence of who-marking) would be false, and would lead to the prediction that attachment preferences should consistently favor the animate NP. Likewise, if we were able to construct matrix clauses $M$ such that explanations involving one NP were more likely to be expressed as a new sentence than explanations involving the other NP, we would again see different predictions.