Anticipatory looks reveal expectations about discourse relations

Citation for published version:
Rohde, H & Horton, WS 2014, 'Anticipatory looks reveal expectations about discourse relations', *Cognition*, vol. 133, no. 3, pp. 667-691. https://doi.org/10.1016/j.cognition.2014.08.012

Digital Object Identifier (DOI):
10.1016/j.cognition.2014.08.012

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Cognition

Anticipatory looks reveal expectations about discourse relations

H. Rohde[*a] & W. S. Horton[b]

[a]University of Edinburgh, Department of Linguistics & English Language, Edinburgh, UK

[b]Northwestern University, Department of Psychology, Evanston, IL, USA

* Corresponding author

Email address: Hannah.Rohde@ed.ac.uk (H. Rohde)

Abstract

Previous research provides evidence for expectation-driven processing within sentences at phonological, lexical, and syntactic levels of linguistic structure. Less well-established is whether comprehenders also anticipate pragmatic relationships between sentences. To address this, we evaluate a unit of discourse structure that comprehenders must infer to hold between sentences in order for a discourse to make sense—the intersentential coherence relation. In a novel eyetracking paradigm, we trained participants to associate particular spatial locations with particular coherence relations. Experiment 1 shows that the subset of listeners who successfully acquired the location~relation mappings during training subsequently looked to these locations during testing in response to a coherence-signaling intersentential connective. Experiment 2 finds that listeners' looks during sentences containing coherence-biasing verbs reveal expectations about upcoming sentence types. This work extends existing research on prediction beyond sentence-internal structure and provides a new methodology for examining the cues that comprehenders use to establish relationships at the discourse level.

1. Introduction

The nature of a coherent discourse is that the utterances within it do not appear together arbitrarily but, rather, relate to each other in meaningful ways. In dialogue, speakers' questions and answers help signal how nearby utterances relate: An utterance may prompt a question (e.g., "Why?" or "What happened next?") and the response is likely to provide the relevant information (about, say, the cause or consequence). Dialogues thus contain frequent overt signals to the relationships that hold between utterances, and speakers track these relationships to understand the structure and content of the conversation. In monologues or single-author texts, intersentential relations are likewise crucial for understanding the content, but comprehenders must often draw their own inferences about what implicit question a particular sentence answers.

The inference of implicit questions or COHERENCE RELATIONS (Asher, 1993; Asher & Lascarides, 2003; Hobbs, 1979; Kehler 2002; Mann & Thompson, 1988; Roberts, 1996) is part of what makes a discourse more than just an arbitrary sequence of sentences. An open question in the experimental pragmatics literature is whether pragmatic inferences are drawn after, during, or even before the relevant sentences are processed in their entirety: Do listeners wait to hear complete propositions in order to initiate discourse-level processing (Garnham, Traxler, Oakhill, & Gernsbacher, 1996; Stewart, Pickering, & Sanford, 2000) or is the identification of pragmatic dependencies incremental, allowing for the integration of cues as they become available (Gibbs, 2002; Koornneef & van Berkum, 2006; Tiemann, Schmid, Bade, Rolke, Hertrich, Ackermann, Knapp, & Beck, 2011; van Berkum, Brown, & Hagoort, 1999), or even anticipatory, reflecting expectations about upcoming dependencies (Kaiser & Trueswell, 2004; Pyykkönen & Järvikivi, 2010)? Existing research has posed this question for a variety of pragmatic phenomena from presupposition and implicature to coreference. Anticipatory pragmatic effects are also relevant to sentence processing more generally, given that pragmatic processes that start mid-sentence can influence within-sentence processing, such as the resolution of local syntactic ambiguity (Altmann & Steedman, 1988; Crain & Steedman, 1985; Desmet, De Baecke, & Brysbaert, 2002; Rohde, Levy, & Kehler, 2011; Sedivy, 2002; van Berkum, Brown, & Hagoort 1999).

In this paper, we ask whether anticipation plays a role in the establishment of coherence relations, and we introduce a novel methodology for measuring such anticipation. Previous work generally has not addressed the time course of coherence establishment, since it is not immediately apparent how a coherence relation could be inferred between a pair of sentences before the content of both sentences is known. Here, we test two cues that have been shown in offline studies to guide preferences about the coherence relation that holds between two sentences, in order to establish whether those cues yield online anticipatory effects regarding the relation between the current sentence and an upcoming sentence. The cues we manipulate involve between-sentence connectives (Simner & Pickering, 2005) and sentence-internal lexical semantics (Kehler, Kertz, Rohde, & Elman, 2008), thereby allowing us to examine how intersentential and intrasentential cues influence online coherence establishment.

2. Intersentential coherence relations

To understand the role of coherence relations in discourse processing, consider the following:

(1) Beryl applauded John. She admired him.

(2) Beryl applauded John. She pleased him.

The first sentences in (1-2) are the same, and the second sentences are structurally and referentially similar—each describes a state of affairs involving two individuals, in both cases re-mentioning the two referents from the first sentence in the same syntactic positions and with the same tense marking on the verb. They differ, however, at semantic and pragmatic levels:

Admiration and pleasure differ in their lexical semantics, and that difference yields a contrast in the relationship between the two sentences in (1) and (2). A natural interpretation of (1) is that the second sentence explains the event in the first sentence, whereas in (2), the second sentence is likely to be understood as describing a consequence of that event. These two relations represent two examples from a larger inventory of relations proposed in the literature (Asher, 1993; Asher & Lascarides, 2003; Hobbs, 1979; Kehler, 2002; Mann & Thompson, 1988).[1]

Note that in (1-2), identifying the intersentential coherence relation requires inferences about other information—either that applause is an action typically performed for someone admirable or that applause is the kind of action likely to please someone. Neither piece of information is stated overtly, yet these passages fail to make sense if that information is not understood or accommodated. Such inferences are ubiquitous in discourse processing, yet are only apparent if we struggle to identify information that would make a passage cohere, as in (3):

(3) Beryl applauded John. She hated him.

As in (1-2), the sentences in (3) are well-formed and easy to understand, but making sense of this passage requires an inference that is hard to reconcile with knowledge of the real world—namely that hatred could plausibly arise from or result in applause. The very fact that readers may start considering contexts in which (3) makes sense (maybe Beryl believes that John is a nervous and private person who cannot bear to be applauded?) is a testament both to readers' expectations that some coherence relation must hold between adjacent sentences and to the ease with which additional inferences arise beyond what is explicitly stated.

---

[1] Across discourse coherence models, all posited inventories distinguish causes and consequences (e.g., Explanation/Result relations in Kehler (2002) or "because"/"as a result" implicit connectives in Prasad, Dinesh, Lee, Miltsakaki, Robaldo, Joshi, & Webber (2008)). In the present experiments we adapt terminology from the papers whose results we build on: *cause/consequence* for Experiment 1 (Simner & Pickering, 2005) and *cause/occasion* for Experiment 2 (Kehler et al. 2008).

Previous work shows that the establishment of intersentential coherence relations is sensitive to a variety of cues, including but not limited to overt connectives (Fraser, 1999; Prasad, Dinesh, Lee, Miltsakaki, Robaldo, Joshi, & Webber, 2008), coreference (Kehler & Rohde, 2014), visual priming (Kaiser, 2012), prosody (Tyler, 2012), verb class (Kehler et al. 2008), and also the preceding coherence relation (Simner & Pickering 2005). In the present work we focus on two of these cues, the preceding coherence relation (Experiment 1) and verb class (Experiment 2), for our investigation of expectation-driven coherence establishment.

Experiment 1 builds on an offline passage-continuation study by Simner and Pickering (2005, Experiment 3), in which the coherence relation in a prompt guided the relation that participants established between the prompt and their own continuation. Examples (4-5) show two of Simner and Pickering's prompts. Unlike examples (1-2), in which the coherence relation between the two propositions was left implicit, the prompts in (4) and (5) contain two propositions joined by an overt connective ("so" and "because," respectively). The second proposition in (4-5) is held constant but the relation between the two propositions varies: In (4), the first proposition gives the reason for the applause (because she admired him, she applauded him); in (5), it describes the consequence (as a result of her applause, she pleased him).

(4) Beryl admired John so she applauded him. ___________

(5) Beryl pleased John because she applauded him. ____________

In these prompts, the connective marks the relation between the first and second propositions; in turn, that relation determines what type of information has been provided and what information is still missing. In (4), the missing information is the consequence (what happened as a result of the applause?), whereas in (5), it is the cause (why did she applaud?). As predicted by Simner and Pickering's SATISFIED GAP HYPOTHESIS, participants wrote more continuations about consequences than causes when the missing information was the consequence (4) than when it was the cause (5). The prompt's coherence relation thus influenced participants' expectations about how the discourse would continue. Offline passage continuations, however, cannot tell us about the time course of these expectations: Do preferences emerge only after a prompt has been processed in its entirety or are available cues used incrementally? Moreover, the passage-continuation task itself presents a challenge for drawing conclusions about comprehension, since the task involves both the comprehension of a prompt and the production of a subsequent sentence.

In the present work, we consider the time course of individuals' generation of discourse expectations as they comprehend a passage. For Experiment 1 we adapt Simner and Pickering's (2005) materials to test listeners' sensitivity to the intersentential connective within the prompt, both as an immediate cue to the material in the second half of the prompt and as an anticipatory cue regarding what missing information the continuation may provide. This study also serves as an initial test of a new methodology for monitoring listeners' online expectations concerning coherence relations. In Experiment 2 we use the same methodology but manipulate the verb, which is likewise predicted to modulate expectations about upcoming relations.

Our goal in both studies is to evaluate the time course of comprehenders' attention to upcoming coherence relations during comprehension. Under a Clausal Integration account (Garnham, Traxler, Oakhill, & Gernsbacher, 1996; Stewart, Pickering, & Sanford, 2000), inferences regarding intersentential pragmatic relationships can only be made after the structural and semantic properties of the two individual sentences have been determined. On the other hand, comprehenders may consider likely relations much earlier in anticipation of encountering an upcoming sentence. Such a result would be in keeping with evidence of expectation-driven processing at other levels of language structure (sounds: Delong, Urbach, & Kutas, 2005; words: Kamide, Altmann, & Haywood, 2003; syntactic categories: Levy, 2008). As we will demonstrate, measurements of comprehenders' anticipatory looks can reveal preferences for particular coherence relations quite soon after coherence-biasing cues.

3. Measuring Anticipation

Anticipating coherence relations involves speculation about an abstract, discourse-level linguistic dependency. Although this type of dependency can be shown to influence other well-studied phenomena (pronoun interpretation: Rohde, Kehler, & Elman, 2006; Rohde, Kehler, & Elman, 2007; Kaiser, 2009; verb-phrase ellipsis: Kehler, 2000; Kehler, 2002; relative-clause attachment: Rohde, Levy, & Kehler, 2011), establishing direct evidence of coherence expectations poses a unique problem. It can be difficult to determine whether comprehenders actually infer an intended relation if the relevant sentences are themselves well-formed and their content creates no processing difficulty. While unexpected coherence relations yield reading-time slowdowns (Traxler, Sanford, Aked, & Moxey, 1997), such work only establishes integration difficulties, not true anticipation. Although some contexts contain overt connectives or rhetorical questions to signal what role an upcoming sentence plays in a discourse, it is possible to encounter a discourse that is coherent but contains no overt coherence markers. In such cases, there may be no word or structure that singlehandedly disambiguates the intended relation. It is challenging, therefore, to determine the time point at which comprehenders successfully understand a coherence relation in context, let alone whether they anticipate such relations.

Existing work on linguistic expectations frequently uses eyetracking within visual-world paradigms (Tanenhaus, Spivey, Eberhard, & Sedivy, 1995; Tanenhaus & Trueswell, 2005). For example, comprehenders' looking behavior shows that they use the selectional restrictions of verbs to anticipate subsequent reference (Altmann & Kamide, 1999): Both *The boy will move the cake* and *The boy will eat the cake* yield looks to a cake within a visual scene, but such looks precede *cake* only with the verb *eat*, presumably because *eat* restricts its direct object to edible entities. Similarly, comprehenders are sensitive to the coreference biases of certain verbs and can use that information to anticipate who will be mentioned next within a clause (Kaiser & Trueswell, 2004) or in an upcoming clause (Pyykkönen & Järvikivi, 2010): A sentence like *The butler frightened the guitarist in the dining room because he…* yields looks to the butler before the causal connective; the butler is typically identified as the causally implicated individual in an event described with the verb *frighten*, and causal continuations are common following such verbs. However, both these and other studies using anticipatory looking (e.g., Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002; Kamide, Scheepers, & Altmann, 2003; Knoeferle, Crocker, Scheepers, & Pickering, 2005) rely on visual scenes to test comprehenders' anticipation of the subsequent mention of particular objects or individuals visible in that scene. The abstract nature of coherence relations poses a different challenge than measuring comprehenders' anticipation of the mention of concrete, imageable referents.

Here, we adapt a paradigm from research on infant category learning, a field that faces analogous challenges in testing individuals' understanding of abstract categories. The present method was inspired by McMurray and Aslin's (2004; Experiment 3) implicit-learning procedure, in which infants were familiarized to a series of colored shapes that passed behind an occluder and reappeared in one of two locations (e.g., red squares to the left, yellow crosses to the right). Later, when presented with a novel form (a red cross), infants' anticipatory looks to one location or the other revealed whether they favored color or shape in classifying the new object and anticipating where it would reappear.

To investigate adults' coherence expectations, we combine a similar occlusion-based display with the presentation of brief recorded passages. Each passage consists of a pair of sentences that are linked via one of two coherence relations; each coherence relation is reliably associated with one of two locations on a visual occluder. Experiment 1 uses passages that express either a *cause* or *consequence*—i.e., the second sentence in the passage provides information about the cause or consequence with respect to material in the first sentence. In the training phase of Experiment 1, we expose participants to specific mappings between passages expressing *causes* or *consequences* and the left or right sides of a visually presented "classifier." For example, participants might be presented with training items that map *causes* to the left and *consequences* to the right. Then, in a test phase we present prompts containing a specific coherence-biasing cue (a connective) and measure participants' anticipatory looks to the *cause/consequence* locations prior to a continuation sentence. Similarly, Experiment 2 uses passages that establish either a *cause* relation as in Experiment 1 or an *occasion* relation, which requires the identification of a sequence of events, and tests how listeners' expectations are influenced by the type of verb present in the first sentence.

If participants can learn the appropriate location~relation mappings and if they use coherence-biasing cues to anticipate upcoming sentence types, we predict that, in Experiment 1, hearing prompts containing *cause*-biasing cues will yield anticipatory looks to the *cause* location of the classifier, whereas hearing prompts containing *consequence*-biasing cues will yield anticipatory looks to the *consequence* location. On the other hand, if coherence establishment is not expectation-driven, then we expect similar looking patterns regardless of the cue in the prompt. Crucially, both experiments are designed to ensure that participants cannot learn a mapping directly between the specific coherence-biasing cues and a visual location. For

Experiment 1, this means that prompts containing the manipulated *cause*- and *consequence*-biasing connectives are followed by equal proportions of *cause* and *consequence* continuations.

4. Experiment 1: Coherence expectations cued by the relation in the preceding context

This experiment uses an implicit-learning procedure to train participants on a mapping between visual locations and *cause*/*consequence* coherence relations. We then examine anticipatory looking to these same locations to test the hypothesis that participants will be sensitive to a prompt's connective and the coherence relation it signals. Following Simner and Pickering (2005), a prompt that focuses on the consequence should generate expectations for a cause in the continuation, whereas a prompt that focuses on the cause should favor consequence continuations. The connective in the prompt—either "so" or "because"—functions as the critical coherence-biasing cue. In an analysis of the patterns of participants' eye fixations while listening to these prompts, we predict an interaction between where participants look (Location: *cause* vs. *consequence*) and when they look (Time: before vs. after the connective). Specifically, following "so" and "because" (Prompt Bias manipulation) comprehenders should fixate the side of the visual classifier associated with the coherence relation they expect to be established between the prompt and the continuation to follow.

4.1 Method

4.1.1 Participants

Twenty-one native-English-speaking undergraduates at Northwestern University participated in the study for course credit. We eliminated data from two participants due to difficulties with calibration, leaving 19 participants for the main analysis.

4.1.2 Materials

Training phase. The training phase used 60 two-sentence passages like those in (6-7). The first sentence served as the prompt and included two propositions joined by an overt connective (e.g., "before" in (6) and "when" in (7)) while the second sentence continued the passage by providing information about the *cause* or *consequence*.

(6) Prompt: The presentation had barely started before the senator fell asleep.

Cause continuation: The topic was the economy, which didn't interest him.

(7) Prompt: Amanda was coming from work when she got into a car accident.

Consequence continuation: The street was closed for several hours afterwards.

The prompt sentences in the training passages always consisted of two propositions joined with a connective taken from a set of 10 possible connectives ("after", "and", "as", "because", "before", "but", "even though", "so", "when", "while"). The continuation provided *cause/consequence* information relative to the second proposition in the prompt: In (6), the fact that the presentation topic was the economy is the cause of the senator's nap, not the cause of the presentation's start; in (7), the road closure follows as a consequence of the accident, not as a consequence of Anne's commute home. All participants heard the same training items, half of which contained *cause* continuations and the other half *consequence* continuations. Training items with "because" and "so" were followed an equal number of times by *cause* and *consequence* continuations to ensure that the training phase did not introduce any associations between these connectives in the prompt and the type of continuation that followed (the anticipation of which is the target of the test phase). Half the continuations began with material to overtly signal the relation (e.g., *cause*:

"The reason for this was", "His rationale was"; *consequence*: "Within two weeks", "As a result").[2]

Post-training assessment. A set of 24 two-sentence passages measured how successfully each participant learned the intended location~relation mapping. Half of these items contained a "because" or "so" connective in the prompt, with *cause* and *consequence* continuations balanced across these items, again ensuring that the experiment itself would not teach participants a mapping between a prompt's connective and the continuation type to follow.

Test phase. The stimuli for the test phase consisted of 10 practice items, 24 target items, and 48 fillers. As in the training phase, all items were two-sentence passages. Examples (8a-b) illustrate the critical manipulation in the prompt while Examples (9a-b) show *cause* and *consequence* continuations. What remains constant across (8a-b) is the content of the second proposition of the prompt (*she applauded him*). Appendix A lists the full set of target items.

(8) Experiment 1 prompts

a. <u>Beryl admired John</u>$_{\text{CAUSE}}$ so she applauded him. [*consequence* bias]

b. <u>Beryl delighted John</u>$_{\text{CONSEQ}}$ because she applauded him. [*cause* bias]

(9) Experiment 1 continuations

a. *cause*: His incredible piano playing had always impressed her.

b. *consequence*: He grinned proudly at the applause.

---

[2] Many Experiment 1 training items were compatible with either a temporal or causal categorization. This reflects the fact that *cause* continuations often describe a preceding event (as in *The tenant called the landlord after she mopped up the water*. <u>*That morning, the sink and bath on the second floor had started leaking again*</u>.) and consequences necessarily move forward in time (see (7)). In order to encourage the inference of Simner and Pickering's *cause/consequence* categories, our items included descriptions of states of affairs that were not inherently temporally ordered (e.g., *Most commuters get stuck in traffic while* <u>*Ann races along the backroads in the countryside*</u> *because* <u>*she is smart enough to avoid the highway*</u>).

The prompts were taken directly from Simner and Pickering's Experiment 3.[3] Half the prompts contained the connective "so" (8a), signaling that the second proposition of the prompt is preceded by its cause (admiration as the cause of applause); the other half contained the connective "because" (8b), signaling that the second proposition in the prompt is preceded by its consequence (delight as the consequence of applause). As noted by Simner and Pickering for cases like (8a), when the prompt provides the cause of the applause, the missing information for comprehenders is the consequence. Accordingly in (8b), given that the prompt provides the consequence of the applause, the missing information going forward is the cause.

Target continuations (9a-b) provided *cause* or *consequence* information relative to the second proposition of the prompt. Participants heard either a *cause* or *consequence* continuation for each item, and continuation type was fully crossed with the connective manipulation in the prompt. If the prompt influences participants' expectations about the type of continuation that will ensue, half the time their expectations were satisfied and half the time they were violated. Note, however, that although the prompt~continuation relation is counterbalanced, the visuals reinforce the location~relation mapping in all phases of the experiment: *Cause* continuations are accompanied by visual feedback in the *cause* location, *consequence* continuations by visual feedback in the *consequence* location.

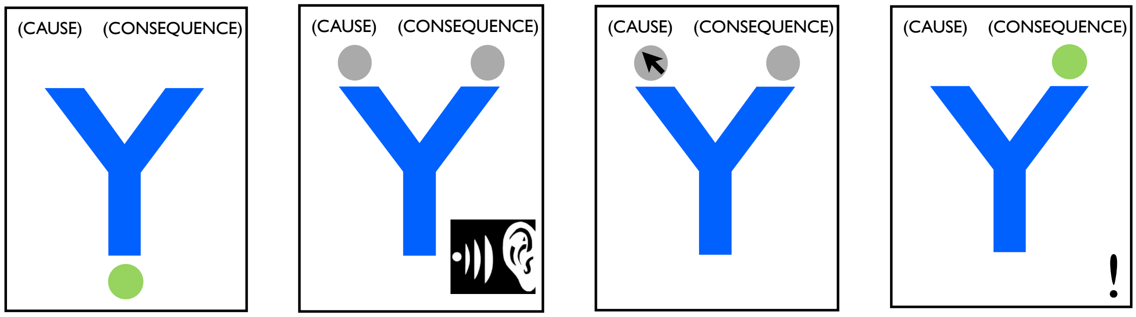
All items were recorded by a male speaker using normal intonation in a sound-attenuating booth at a 22.05 kHz sampling rate. The prompt and continuation for each item were recorded separately, reducing the likelihood that intonational cues in the prompt would signal the upcoming continuation. From these, we prepared individual sound files for use during the training, assessment, and test phases. For all items, we used two sound files, one containing the

[3] Of Simner and Pickering's 26 items, 24 are used here. The two we excluded differed from the others (a negated verb in one case and a verb of physical perception in the other).

prompt and one containing the continuation. Analyses were carried out with respect to the offset of the connective "so" or the end of the first syllable of "because"—chosen as the point at which listeners have specific evidence of the connective linking the prompt's two propositions. The average time from item onset to the offset of the first syllable of "because" ($M = 1269$ ms) was significantly shorter than for the "so" connective ($M = 1438$ ms; $t(43) = 2.63$, $p < 0.05$), due to differences in the length of the first clauses in the two conditions.

We also created a visual display for the "classifier tube," modeled on the one used by McMurray and Aslin (2004). This classifier consisted of a large blue "Y" on a white background (see Figure 1). The same classifier was presented during training, assessment, and test phases. We created separate versions of the display that showed a green "ball" at the bottom of this classifier, the ball coming out of the upper left branch, and the ball coming out of the upper right branch. These additional displays were presented before and after the basic display to give the impression that the ball was traveling through this virtual tube.

Training phase. The training phase was intended to teach participants one of two mappings between specific continuation types and specific spatial locations on the "tube." Figure 1 assigns the left output end to *cause* continuations and the right output end to *consequence* continuations. Mappings were counterbalanced across participants. The experimenter told participants that their task was to figure out how this virtual tube classified a series of two-sentence stories. Participants were given no information about the basis for classification. They were simply told that they would be informed whether each response was correct or not.



a. Initial state
b. Hear prompt and continuation
c. Select left/right output
d. Feedback with ball at one output, plus sound effect

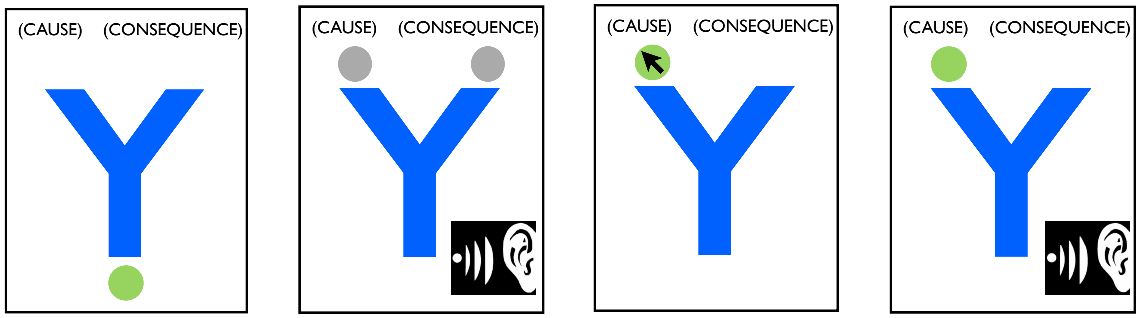
Figure 1. Experiment 1 training phase (category labels were not visible to participants)

4.1.3 Procedure

Training phase. The classifier's initial state for each training trial is shown in Figure 1a. For each item, participants clicked on the ball to "launch" it into the tube. This started the auditory presentation of a recorded passage (1b). Participants were instructed to listen carefully to the passage and to then click on the output end where they thought the ball would re-appear, based on their guess about how the tube was classifying stories (1c). They were told that the two ends of the Y-shaped tube corresponded to two different classifications. The location of the re-emerging ball and a sound effect then provided feedback about whether their response was correct or incorrect (1d).

Post-training assessment. The task was identical to the training but without feedback.

Test phase. Participants were informed that the classifier tube and the corresponding categories were exactly the same as before. They were told that they would again be hearing two-sentence passages, but that the ball would emerge from the tube before the continuation played, and that their task was to listen carefully to the first sentence in each item and then prepare to click on the ball as soon as it emerged from the tube, which would trigger playback of the second sentence.



a. Initial state b. Hear prompt plus 1500 ms pause c. Speeded click d. Hear continuation

Figure 2. Experiment 1 test phase (category labels were not visible to participants)

To start each test trial, participants clicked on the green ball (see Figure 2a) to "launch" it into the bottom of the tube and to initiate playback of the prompt. As participants listened to the prompt (2b), we recorded their point of gaze continuously from the onset of the sound file until the final re-appearance of the green ball. The onset of the prompt also coincided with the appearance of a pair of gray circles, one at the *cause* output end and one at the *consequence* output end. The gray circles served as visual cues for where the green ball could reemerge, and were present throughout the presentation of the prompt and a subsequent 1500 ms pause. After this pause, the gray circles disappeared, and the green ball appeared at either the *cause* or *consequence* location. Once the green ball re-emerged, participants clicked on it (2c) to hear the continuation (2d). The location of the green ball at the left or right output end signaled the upcoming continuation type (e.g., in Figure 2c, when the ball re-appeared on the left, participants always heard a *cause* continuation). Recall, though, that in order to prevent participants from learning a mapping between the coherence-biasing cue and the continuation type, the continuation type indicated by the ball's final location violated the prompt bias on half the target trials.

Prior to the test phase, participants were introduced to the eyetracking equipment, which consisted of a table-mounted infrared camera positioned below the computer monitor, connected to an ASL 6000 eye tracker recording at a sampling rate of 60 Hz. After a brief calibration procedure, participants then began the test phase. Between trials, participants fixated on a single centrally-located cross. One-fourth of the items were followed by a comprehension question to encourage participants to pay attention to the passages. Participants responded to these questions with a mouse click in a “yes” or “no” box on the screen. Following a brief practice block, the 72 test items were presented, interrupted with two brief pauses for re-calibration if necessary. After the test phase was complete, participants were asked to complete a questionnaire regarding the purpose of the experiment and the categories they had inferred.

4.2 Results

4.2.1 Implicit learning

Very few participants (n = 5 out of 19, or 26%) were above chance on the post-training assessment. Chance was determined using an estimate of the probability of seeing the observed data under a binomial distribution, with a cutoff of $p < 0.05$; above-chance performance required 17 correct answers out of 24. Comments from the post-experiment questionnaire show that most participants (including several above-chance participants) could not articulate the classifier categories. Frequent responses included “no idea,” “male/female?” and “positive/negative?” A small number of participants did successfully identify one category that “explains,” “tells the cause,” or “could use *because*” and another category for “what happened after” or the “result.” Below, we report results separately for the above- and at-chance participants.[4]

---

[4] It is a reasonable concern that only five of our 19 participants showed evidence of successful location~relation learning following the training phase. As we discuss later, the complex nature of the materials may have prevented some participants from acquiring the relevant categories; additional training items likely would have been needed to increase the number of participants showing better than chance performance. In the study by McMurray and Aslin (2004, Experiment 3) that inspired the present paradigm, only 9 of 22 infants showed above-chance patterns of looking on training trials. Rather than presenting data from above-chance participants

4.2.2 Eye gazes during the test phase

As an index of participants' ability to learn the two coherence categories and to use cues in the prompt to anticipate the continuation type, we were interested in participants' eye gazes to the output ends of the classifier during presentation of the prompt and before the reappearance of the ball. We defined two critical areas of interest (AOI) on the display, consisting of two 250 × 150 pixel rectangles centered on the middle of the output ends of each arm of the classifier. Each gaze sample was then coded as landing either within one of these two regions or in neither region. Depending on the specific location~relation mapping presented to the participant during the training phase, one AOI (either left or right) was designated as the *cause* location while the other AOI was designated the *consequence* location. Coding began at the onset of the prompt and ended as the ball reappeared at one arm of the tube, signaling playback of the continuation.

To examine how gazes to these locations were synchronized with the accompanying speech, we identified the offset of the first syllable of the critical connective within each prompt (henceforth 'connective offset'). All our analyses are aligned to this offset, which establishes a consistent starting point for the critical post-connective region. In our analyses, we examined the time window starting 1000 ms before this offset, since this includes the verb in the first half of the prompt (average time between verb offset and connective offset was 548 ms, with a maximum time of 979 ms). We analyze eye data up to 1500 ms after connective offset, which includes the rest of the prompt (average time between connective offset and prompt offset was 925 ms, with a maximum time of 1550 ms). Note that all relevant eye fixations occurred during playback of the prompt, and there was no variation by condition in the 1500ms pause.

---

alone, we report the results from both above-chance and at-chance participants separately, in keeping with previous classification learning studies that use training phase performance to divide participants into "learners" and "non-learners" when examining data from a subsequent test phase (e.g., Filoteo et al., 2005; Weiler, Bellebaum, and Daum, 2008).

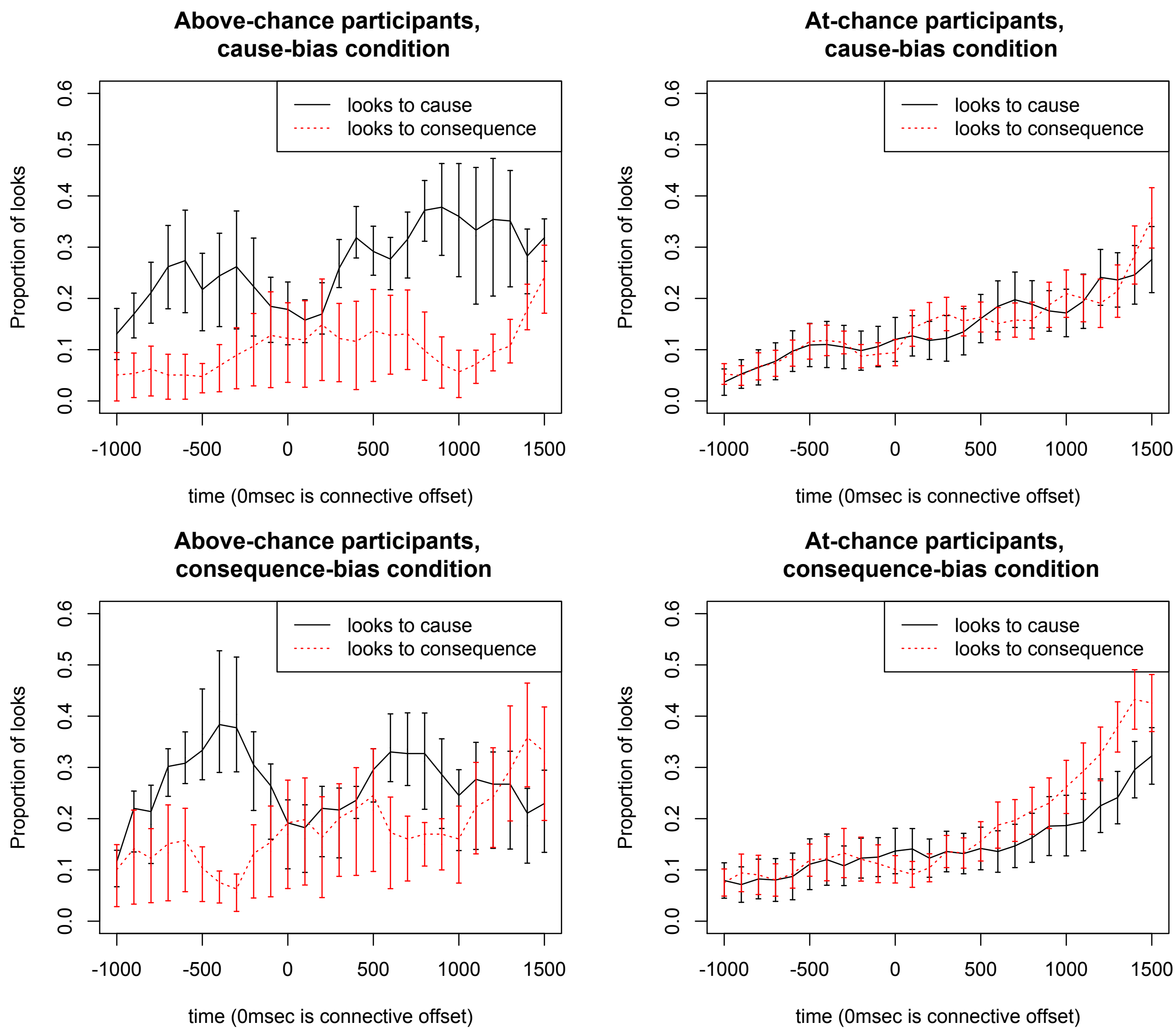


Figure 3: Experiment 1 results for above-chance and at-chance participants. Error bars show standard error of the mean for each 100 ms bin.

Figure 3 presents an overview of the time course of the mean proportions of looks to each screen location for both Prompt Bias conditions, computed in discrete 100 ms bins centered at connective onset. Separate plots are given for above-chance and at-chance participants. For the above-chance participants, these plots reveal an early tendency in both bias conditions to look to the *cause* location before the connective. After the connective, this tendency to look to the *cause* location strengthens in the *cause*-bias condition but becomes weaker in the *consequence*-bias condition. For the at-chance participants, there are no strong tendencies to look differentially at either screen location, in either bias condition. To analyze these patterns, we calculated the overall proportions of looks to the *cause* and *consequence* locations for two periods of time during presentation of the prompt: before connective offset and after connective offset. Table 1 shows the average proportions of looks to each location out of all gaze samples, broken down by prompt bias, time window, and training performance.

Table 1. Proportions of looks in Experiment 1 to the *cause* and *consequence* locations by prompt bias, time window, and participants' performance on the post-training assessment

		Pre-connective window		Post-connective window	
Training Performance	Prompt bias	*cause* Location	*consequence* Location	*cause* Location	*consequence* Location
Above chance	*cause* bias	0.22	0.07	0.29	0.12
	consequence bias	0.28	0.12	0.26	0.22
At chance	*cause* bias	0.09	0.09	0.18	0.18
	consequence bias	0.10	0.10	0.18	0.22

To examine whether participants used the first half of the prompt and the intersentential connective to anticipate upcoming material, we created two parallel sets of mixed-effect models, one set for the above-chance participants and one for the at-chance participants. We modeled looking behavior with three factors (and their interactions): Prompt Bias (*cause*-bias vs. *consequence*-bias), Location (*cause* vs. *consequence*), and Time (pre-connective vs. post-connective). Because our dependent measure is categorical—that is, we are interested in which screen locations participants were fixating as they listened to the auditory sentences—our analyses use weighted empirical logit regression, an approximation to multilevel logistic regression (McCullagh & Nelder, 1989) that involves using an empirical logit function to transform our data, placing them on a log odds scale appropriate for categorical observations (Barr, 2008; Jaeger, 2008).[5] Additionally, as described by Barr (2008), the nonindependence of gaze fixations necessitates the aggregation of trials within a given condition and the grouping of observations within discrete temporal windows when using empirical logit regression. Because this aggregation procedure must be carried out separately across participants and across items, we conducted separate by-participant and by-item analyses.

Following these procedures, we fit our transformed data for the above-chance and at-chance participants to mixed-effect models that included Prompt Bias, Location, and Time and their interactions as fixed effects. These models included either participant or items as a random intercept, as appropriate. All models we report in this paper also included, whenever possible, by-participant or by-item random slopes for fixed factors, as well as their interactions (Barr, Levy, Scheepers, & Tily, 2013). If this maximal model failed to converge, the highest-order slope associated with the least amount of variance was progressively removed until a model was found that successfully converged. Fixed effects were mean-centered to reduce collinearity, and all models were fit using the *lmer* function in the `lmer4` package in R (Bates, Maechler, & Bolker, 2011) using maximum-likelihood estimation. All p-values reflect model comparisons based on a likelihood-ratio $\chi^2(1)$ test of the change in the goodness of fit between the full model and a comparison model with the relevant fixed effect removed.

[5] We continue to report raw proportions in tables and figures for clarity.

Table 2. Experiment 1 coefficients and standard errors of the estimates for Location, Prompt Bias, and Time, by participants' performance on the post-training assessment

	Above-Chance (N = 5)				At-Chance (N = 14)			
	Participants		Items		Participants		Items	
Fixed Effect	β	SE	β	SE	β	SE	β	SE
Location	-0.94**	0.07	-0.55**	0.11	0.32	0.16	0.07	0.05
Prompt Bias	0.36*	0.07	0.23	0.11	0.18†	0.05	0.16†	0.05
Time	0.56*	0.07	0.16*	0.07	1.11**	0.06	0.77**	0.04
Location × Prompt Bias	0.52†	0.14	0.22	0.23	0.19	0.10	0.14	0.11
Location × Time	0.82*	0.17	0.26†	0.14	-0.09	0.15	0.28*	0.09
Prompt Bias × Time	-0.40	0.14	-0.32*	0.13	-0.40	0.10	0.00	0.09
Location × Prompt Bias × Time	0.72	0.28	0.70†	0.34	0.39	0.19	0.41	0.21

Note. P-values were obtained by a model-comparison approach based on a likelihood-ratio χ2(1) test of the change in the goodness of fit between the full model and a comparison model in which only the relevant fixed effect was removed.
† p < .10 * p < .05 ** p < .01

Table 2 reports the coefficients and standard errors of the estimates for each of the fixed effects in the above-chance and at-chance models. First, for the above-chance participants, there were main effects of Location and Time—more looks to the *cause* location than *consequence* and more looks in the post-connective window than the pre-connective window. In addition, several effects were significant in either the participants or items analysis but not both. These included a trend towards a main effect of Prompt Bias, with more overall looks in the *consequence*-bias condition than the *cause*-bias condition; this is possibly due to the later arrival of the *consequence* cue, the connective "so," which would have permitted more time for participants to look to either of the two locations. This pattern was driven by a trend towards a Prompt Bias × Time interaction, whereby the higher proportion of looks in the *consequence*-bias prompts was more apparent in the pre-connective window, in keeping with the later arrival of "so" than "because". A marginal Location × Time interaction shows the main effect of Location favoring the *cause* location more in the pre-connective window than the post-connective window. Finally, there was a trend towards a 3-way interaction between Location, Time and Prompt Bias, whereby the overall preference to look to the *cause* location was reduced after the connective in the *consequence*-bias condition but not in the *cause*-bias condition, although this was not reliable in either analysis.

Of primary interest were the follow-up comparisons to test for two-way Location × Time interactions; these analyses test whether the above-chance participants' pattern of looking changed over time in either the *consequence*-bias or *cause*-bias conditions. The interaction is significant for the *consequence*-bias condition (β_{subj}=1.17, *SE*=0.18, $p<0.01$; β_{item}=0.59, *SE*=0.11, $p<0.05$), with an increase in looks to the *consequence* location after the connective, but not for the *cause*-bias condition (β_{subj}=0.42, *SE*=0.27, p=0.49; β_{item}= -0.10, *SE*=0.15, p=0.72), in which the bias towards the *cause* location remains consistent before and after the connective. The pattern of results is in keeping with the claim that the coherence relation in the prompt induced a shift in participants' looking patterns: The overall *cause* preference was maintained following the *cause*-biasing connective "because" and was reduced following the *consequence*-biasing connective "so."

For the at-chance participants, Time was the only reliable effect, with a higher proportion of looks after the connective than before. There was a marginal effect of Prompt Bias, favoring the *consequence*-bias prompts. This matches the Prompt Bias trend observed with the above-chance participants, again possibly due to the later arrival of the *consequence*-bias cue compared to the *cause*-bias cue. No other effects or interactions reached significance.

In sum, then, these analyses showed higher proportions of looks following the connective than before the connective, for participants who had performed both above and at chance on the post-training assessment. For the above-chance participants, the most reliable effect was a significant preference to look toward the *cause* location, a preference that was reduced following the connective in the *consequence*-bias condition due to an increase in looks to the *consequence* location. This overall preference to look to the *cause* location is somewhat unexpected given the consequence bias reported by Simner and Pickering (2005), but an examination of the materials reveals that the majority of the verbs used in first half of the prompts were verbs marking implicit causality, such as *applaud* and *please*. We return to this point in the Discussion, where we motivate a more systematic examination of verb class in Experiment 2.

Despite this early preference for the *cause* location, the patterns of fixations after the connective follow a pattern that is consistent with the bias introduced by the connective, at least for the above-chance participants. As shown in Table 1, the early preference for the *cause* location continued after the *cause*-biasing "because," even strengthening slightly. In contrast, the initial *cause* preference weakened considerably after the *consequence*-biasing "so," with a substantial proportion of fixations to the *consequence* location. Although this effect did not flip entirely in favor of the *consequence* location, the overall pattern does suggest that participants who showed evidence of learning the classification categories during training were sensitive to the prompt manipulation. From these analyses, we have evidence to support the claim that listeners, after having learned to associate specific screen locations with specific coherence relations, are more likely to look to those same locations depending on a coherence-biasing cue in the prompt.

4.3 Discussion

The training results show that participants struggled to learn the classifier's mapping between coherence relations and visual locations. This may reflect the complexity of the prompt sentence with its two propositions and the fact that the classification categories reflected the relationship between second proposition and the continuation. Nonetheless, the results from the test phase provide a picture of participants' early coherence biases and the way those biases are modulated by the connectives "so" and "because." Participants who did learn the location~relation mapping showed an overall preference to look to the *cause* location over the *consequence* location—a preference which appeared early, during the first half of the prompt. This early *cause* preference was maintained following the *cause*-biasing connective and was reduced following the *consequence*-biasing connective. This pattern emerges as a significant Location × Time interaction for the *consequence*-biasing prompts and no significant interaction for the *cause*-biasing prompts. These patterns occur before the end of the prompt, indicating that participants were responding to the connective without waiting for the entirety of the two propositions in the prompt.

There are two interpretations, however, for the differential looking following the "because" and "so" connectives. On one hand, it is compatible with Simner and Pickering's Satisfied Gap Hypothesis: The prompts differed in what type of information was missing, and above-chance participants looked more to the *consequence* location (and less to the *cause* location) when the missing information was the consequence than when the missing information was the cause. On the other hand, the same looking behavior could merely reflect participants' classification of the second half of the prompt: The connective "so" signals that the second proposition in the prompt describes a consequence, and above-chance participants looked more to the *consequence* location when the connective marked a consequence ("so") than a cause ("because"). Only the first interpretation provides support for coherence anticipation. Note, however, that under either interpretation, the above-chance participants are responding to the connective before they know the full content of the two propositions in the prompt. Experiment 2, however, will provide evidence that is most compatible with an anticipation account.

As indicated previously, the prevalence of implicit causality (IC) verbs in the first half of the prompt may have induced a bias towards a cause. IC verbs guide inferences about causality and coreference (Au 1986; Brown & Fish, 1983; Ferstl, Garnham, & Manouilidou, 2011; Garvey & Caramazza, 1974; Koornneef & Van Berkum, 2006; McKoon, Greene, & Ratcliff, 1993; Pyykkönen & Järvikivi, 2010; Stevenson, Crawley, & Kleinman, 1994) as well as coherence (Kehler, Kertz, Rohde, & Elman, 2008). In a series of passage-continuation studies, Kehler et al. (2008) found that context sentences with different classes of verbs favor different coherence relations. Continuations following prompts with IC verbs are more likely to provide an explanation of the IC eventuality. Thus, the prevalence of IC verbs in the present materials may have driven the above-chance participants' early expectations that the discourse would continue in a manner that provided information about the cause. Experiment 2 therefore will test this verb-driven coherence bias by comparing looks following verbs with opposing coherence biases. In particular, we contrast IC verbs with transfer-of-possession verbs, which Kehler et al. (2008) showed favor continuations that describe what happened next.

Additionally, Experiment 2 eliminates the biclausal prompts. Although some Experiment 1 participants learned the location~relation mapping, many struggled in training to figure out the abstract categories and use those categories during the test phase. Experiment 2 will use simpler contexts so that participants need only infer that the relevant categories are the relations between a monoclausal prompt and a monoclausal continuation (as opposed to tracking the relationship between the second proposition of a biclausal prompt and a continuation). This is expected to increase the number of participants who perform above chance on the post-training assessment.

Despite these caveats, it is important to point out that Experiment 1 demonstrates the effectiveness of our novel learning-and-looking methodology: Participants who successfully learn the appropriate location~relation mappings are able to use their knowledge of those relations quite early. In particular, above-chance participants' preferential looking toward the *cause* location prior to the connective is noteworthy because it indicates that participants quite readily look to a relevant location even during time windows when the task does not specifically encourage anticipation. This suggests that introducing a coherence-biasing cue early in a sentence could yield preferential looking to the location associated with the anticipated coherence relation. We test this in Experiment 2.

5. Experiment 2: Coherence expectations cued by the verb of the current sentence

This experiment tests the hypothesis that participants' coherence expectations are sensitive to verb class. This hypothesis stems from Kehler et al.'s (2008) offline story-continuation results showing that the class of verb in a prompt affects the type of continuations participants produce. Kehler et al. discuss the effect of verb class on a variety of coherence relations, two of which are relevant to the current study: Explanation and Occasion relations. Specifically, prompts with IC verbs (e.g., *admire, please, scold*) were shown to yield a bias for continuations that explain the eventuality described in the prompt, whereas transfer-of-possession (TOP) verbs (e.g., *hand, give, ship*), yielded a bias for continuations that describe a subsequent event occasioned by the eventuality in the prompt. For our purposes, we will refer to those two relations as *cause* and *occasion*. The inference of an *occasion* relation requires the identification of a pair of events in which the end state of one serves the start state of the next (Kehler, 2002; see also Hobbs, 1990). For example, a passage like "Elizabeth delivered a letter to Darrell. He opened it with curiosity." sets up an event in the first sentence whose conclusion (the delivery of the letter) is the onset of the event in the next sentence (the opening of the letter). By that definition, the *occasion* category subsumes the *consequence* category of Experiment 1. Whereas *consequences* require the inference of a causal connection between two juxtaposed events (e.g., "The cyclist broke his leg. He was in a cast for months."), *occasions* only constrain the relationship between the end state of one event and the start state of the next such that the passage conveys the forward progression of time (e.g., "The cyclist took his bike out of the garage. He pedaled down the street.").

If the verb-driven biases observed in offline studies influence online processing (as suggested by the early *cause* preference in Experiment 1) and if participants can learn a location~relation mapping, we predict divergent looking preferences by verb class. The classifier tube for this experiment will distinguish two continuation types: *cause* and *occasion*. In the analysis of participants' eye data, we predict an interaction between where they look (Location: *cause* vs. *occasion*) and when they look (Time: before vs. after the verb) for both coherence-biasing cues (Verb Class: IC vs. TOP). We predict a 3-way interaction whereby location preferences emerge only after the verb —IC verbs yielding a *cause* bias; TOP verbs yielding an *occasion* bias. Additionally, we also examine specific time windows (100 ms bins before and after the verb) to trace the effects' time course (as represented by interactions between Location and Verb Class in particular temporal bins).

5.1 Method

5.1.1 Participants

Forty-nine native-English-speaking Northwestern undergraduates participated in the study for course credit. We removed data from seven participants due to calibration difficulties or a failure to follow task instructions, leaving 42 participants for the main analysis.

5.1.2 Materials

Training phase. The initial training phase used 60 two-sentence passages like those in (10-11):

(10) Prompt: Leo takes the bus to work.

Cause continuation: He doesn't have a car.

(11) Prompt: Amanda was in a car accident.

Occasion continuation: The street was closed for several hours afterwards.

Half the training items contained a *cause* continuation (10) and half contained an *occasion* continuation (11). None of the training items contained IC or TOP verbs.[6]

Post-training assessment. A set of 24 two-sentence passages measured training success. Half of these items contained an IC or TOP verb in the prompt, with *cause* and *occasion* continuations

[6] As in Experiment 1, some items were compatible with the inference of both temporal and causal categories. Whereas Experiment 1 builds on the *cause/consequence* categories from Simner and Pickering (2005), Experiment 2 targets a *cause/occasion* contrast based on Kehler et al.'s (2008) finding that IC and TOP verbs differ most in the proportions of Explanation and Occasion continuations they yield. Although *occasions* move forward in time and may optionally permit an additional causal connection, *consequences* require a causal inference. To determine whether the training phases of Experiments 1 and 2 reflected this difference, we conducted a norming study on the training items from both experiments. Participants (N=18) were recruited from Amazon Mechanical Turk, and for each passage they were asked to indicate on a 1-7 scale "How likely is it that the second sentence describes something that not only happened AFTER the first sentence (rating=1) but also arose AS A RESULT of the first sentence (rating=7)?" Although participants generally favored causal interpretations (M=4.96), the results confirmed that the Experiment 2 training items favored the inference of a causally-defined category (M=4.46) less than the Experiment 1 training items (M=5.47; $p<0.05$ via a likelihood-ratio test).

balanced across these items to prevent participants from learning a mapping between verb class and continuation type (the anticipation of which is the target of the test phase).

Test phase. The stimuli for the test phase consisted of 10 practice items, 80 target items, and 80 fillers. For the target items (see examples (12) and (13)), the verb class in the prompt was manipulated within participants and between items. Half of the target items contained IC verbs while the other half contained TOP verbs. The complete list of target items can be found in Appendix B. Participants heard either the *cause* or *occasion* continuation for each item, and the choice of continuation type was again crossed with the verb class manipulation to ensure that nothing in the experimental design would link the presence of a particular verb class in the prompt with the upcoming continuation type. If participants use the verb to anticipate the upcoming continuation, their expectations will be violated half the time.

(12) IC prompt: Arthur scolded Patricia in the hallway. [*cause* bias]

a. *cause* continuation: She had put thumbtacks on the teacher's chair.

b. *occasion* continuation: He then sent her to the principal's office.

(13) TOP prompt: Heidi shipped Eric a package. [*occasion* bias]

a. *cause* continuation: She thought he'd like some cookies from home.

b. *occasion* continuation: He wrote her a thank you note.

IC verbs were taken from Kehler et al. (2008), which in turn were adapted from McKoon et al. (1993). The TOP verbs were drawn from Stevenson et al. (1994) and Ferretti, Rohde, Kehler, & Crutchley (2009). IC verbs with Experiencer/Stimulus or Agent/Patient arguments appeared in a subject-verb-object frame (*Arthur scolded Patricia*). For transfer events, the Source/Theme/Goal arguments can be expressed either in a double-object frame (*Heidi shipped Eric a package*) or with a *to*- prepositional phrase (*Heidi shipped a package to Eric*). The studies from which the

TOP verbs were adopted used only the prepositional-phrase frame, but we included the double-object frame in a quarter of our TOP items in order to better mask the verb class manipulation.

On average, the verb in the prompt ended 956 ms following sentence onset and the prompts ended from 247 ms to 1558 ms (M = 830 ms) following the verb. The average offset for IC verbs (M = 1147 ms) occurred significantly later than the offset for TOP verbs (M = 765 ms; $t(78) = 4.60, p < .001$), due to differences in the length of the verbs themselves (IC verbs tended to be multisyllabic whereas TOP verbs tended to be monosyllabic) and to the presence of several sentence-initial prepositional phrases in the IC prompts. These phrases were included in half of the IC items either sentence-initially or sentence-finally in order to provide more context for the eventuality evoked by the short IC sentences. These prepositional-phrase adjuncts also served to reduce the difference between the TOP prompts with three arguments (Source/Theme/Goal) and the IC prompts with only two arguments (Stimulus/Experiencer or Agent/Patient).

Both the *cause* and *occasion* continuations were constructed to be sensible continuations to the prompt. Each participant saw only one version of each target item, and continuation type was counterbalanced: IC prompts were followed half the time with a *cause* continuation and half the time with an *occasion* continuation; likewise for TOP prompts. Practice and filler items were similar to the training items and contained no IC or TOP verbs. Fillers contained equal numbers of *cause* and *occasion* continuations. All *cause* and *occasion* continuations were visually reinforced with the ball appearing at the associated screen location, as in Experiment 1.

All of the items were recorded by a male speaker using normal intonation in a sound-attenuating booth at a 22.05 kHz sampling rate. Prompts and continuations were always recorded separately. From these recordings, we prepared individual sound files for each item for use during the training, assessment, and test phases of the study. Experiment 2 used the same "classifier tube" display as Experiment 1 (see Figure 2).

5.1.3 Procedure

Training phase. As in Experiment 1, the training phase was intended to teach participants one of two mappings between specific coherence relations and specific spatial locations on the visual classifier. The two possible location~relation mappings were counterbalanced across participants, and participants were given no information about the basis for classification. The training procedure matched that used in Experiment 1; participants listened to the pre-recorded sentences and attempted to identify the relevant classification by clicking on one of the two output ends of classifier, receiving feedback about the accuracy of their decision. Training continued until the participant either made 10 correct selections in a row or listened to all training trials. Fifteen participants (36%) ended training early via this criterion.

Post-training assessment. Again, the post-training assessment was identical to the training phase except there was no feedback. Regardless of their performance on these assessment trials, all participants immediately continued to the test phase.

Test phase. The test phase trials followed the same procedures as Experiment 1. Following a brief practice block, the 160 items in the test phase were presented in blocks of 20 to 32 items, with a brief pause between blocks for re-calibration if necessary. After the test phase was complete, participants were asked to complete a questionnaire regarding the purpose of the experiment and the categories they had inferred.

5.2 Results

5.2.1 Implicit learning

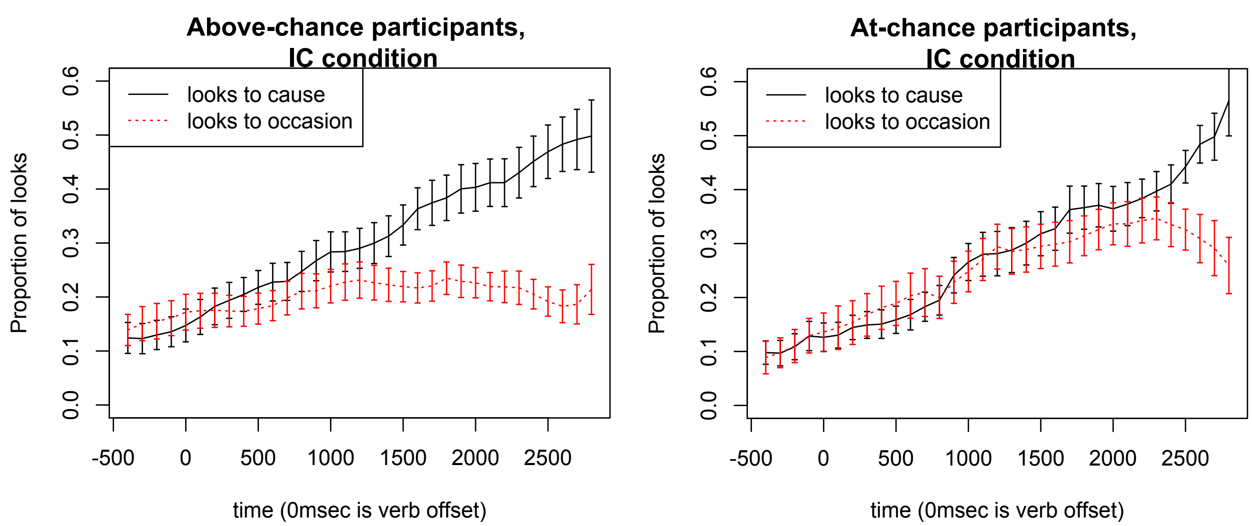
The post-training assessment revealed that a little over half the participants (n = 26 out of 42, or 62%) were above chance in guessing the category of the post-training passages. As in Experiment 1, comments from the post-experiment questionnaire suggest that most participants (including 12 above-chance participants) could not articulate the classification categories.

5.2.2 Eye gazes during the test phase

As an index of participants' ability to learn the two coherence categories and to use the verb class from the initial sentence to anticipate the continuation type, we analyzed participants' eye gazes to the relevant screen locations during presentation of the prompt and before the reappearance of the ball. AOI measurements and gaze coding followed that in Experiment 1.

To examine how gazes to these locations reflect the accompanying speech, all our analyses are synchronized to the offset of the verb in each prompt. We consider the time window from 400 ms before verb offset to 2800 ms after verb offset, which includes the remainder of the prompt sentence plus the 1500 ms post-prompt pause that occurred before re-emergence of the ball and playback of the continuation sentence. Unlike Experiment 1, in which we truncated analysis at the end of the prompt, here we included the post-prompt pause because some relevant eye fixations, particularly for the at-chance participants, emerged after playback of the prompt.

Figure 4 shows the time course of the mean proportions of looks to each screen location for both verb classes. Separate plots are given for above-chance and at-chance participants. For both groups, these plots reveal patterns of looking that are in keeping with our predictions: Participants looked more to the *cause* location following IC verbs and more to the *occasion* location following TOP verbs. The effects emerge earlier for the above-chance participants. To analyze these patterns, we first calculated the overall proportions of looks to the *cause* and *occasion* locations for the time windows before and after verb offset. Table 3 shows average proportions of all gaze samples within each window (including looks to neither AOI), broken down by time window, verb class, and training performance.



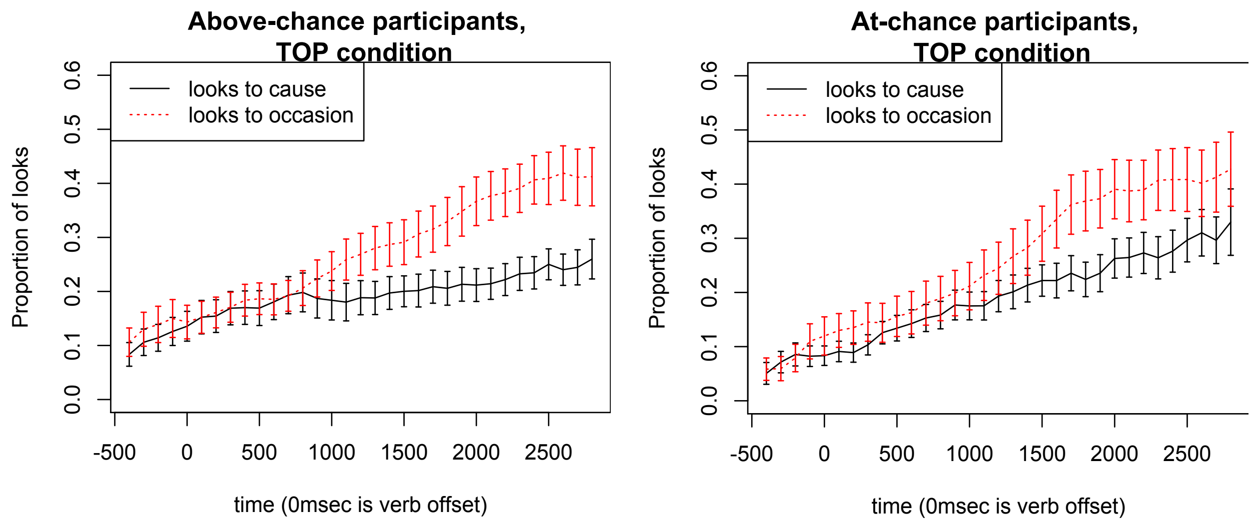


Figure 4: Experiment 2 results for above-and at-chance participants, across IC and TOP verbs

Table 3. Proportions of looks in Experiment 2 to the *cause* and *occasion* locations by verb class, time window, and participants' performance on the post-training assessment

		Pre-Verb window		Post-Verb window	
Training Performance	Verb Class	*cause* Location	*occasion* Location	*cause* Location	*occasion* Location
Above chance	IC Verb	0.15	0.18	0.53	0.17
	TOP Verb	0.11	0.14	0.20	0.42
At chance	IC Verb	0.11	0.10	0.42	0.22
	TOP Verb	0.07	0.08	0.20	0.36

The overall proportions of looks suggest quite strongly that participants were generating expectations about the upcoming discourse continuation based on the verb in the prompt. Specifically, in the post-verb window, IC verbs were accompanied by a greater proportion of looks to the *cause* location than the *occasion* location. Conversely, TOP verbs were followed by a greater proportion of looks to the *occasion* location than the *cause* location. These patterns are evident both for participants who performed above chance and those at chance on the post-training assessment. To examine these patterns, we created two parallel sets of mixed-effect models, one for the above-chance participants and one for the at-chance participants. We tested how looking behavior was influenced by three factors (and their interactions): Location (*cause* vs. *occasion*), Verb Class (IC vs. TOP), and Time (pre-verb vs. post-verb). As described in the Results section of Experiment 1, our analyses use weighted empirical logit regression. Again, the aggregation process necessitates separate by-participants and by-items analyses.

Table 4. Experiment 2 coefficients and standard errors of the estimates for Location, Verb Class, and Time, by participants' performance on the post-training assessment

	Above-Chance (N = 26)				At-Chance (N = 16)			
	Participants		Items		Participants		Items	
Fixed Effect	β	SE	β	SE	β	SE	β	SE
Location	-0.17	0.06	0.03	0.01	0.05	0.05	0.21**	0.03
Verb Class	-0.06	0.01	-0.09**	0.01	-0.17**	0.01	-0.14**	0.01
Time	1.05**	0.03	0.79**	0.01	1.47†	0.05	1.22**	0.02
Location × Verb	0.87	0.04	0.87**	0.02	0.44	0.07	0.52**	0.05
Location × Time	-0.19	0.06	-0.20**	0.02	0.27	0.06	0.12	0.04
Verb × Time	0.06	0.02	0.06	0.02	0.22	0.04	0.12	0.04
Location × Verb × Time	1.29**	0.07	0.97**	0.04	0.36	0.11	0.47*	0.09

Note. P-values were obtained via the same model-comparison approach used in Experiment 1.

† $p < .10$ * $p < .05$ ** $p < .01$

Table 4 reports the coefficients and standard errors of the estimates for each of the fixed effects in the models, computed for the above-chance and at-chance data separately. First, for the above-chance participants, there was a main effect of Time due to more looks in the post-verb time window than the pre-verb time window. In addition, we find the predicted Location × Verb Class × Time interaction due to looking preferences that emerged after the verb and differed by verb class. The remaining trends were significant in either the participants or items analysis but not both. These include a trend towards a main effect of Verb Class, with more looks in the IC condition than the TOP condition, and a trend towards a Verb Class × Location interaction, likely driven by the strong interaction of Verb Class and Location after the verb. There was also a trend towards a Location × Time interaction, whereby the bias to the *cause* location was stronger after the verb; this was likely driven by the Location × Verb Class × Time interaction. The remaining effects and interactions did not reach significance.

Of primary interest are the follow-up comparisons to test for the critical two-way Location × Time interactions for the IC and TOP sentences, since those analyses indicate whether participants' pattern of looking changed in response to the coherence-biasing cues. Again, focusing first on the above-chance participants, the interaction is significant for the IC condition (β_{subj}= -0.79, SE=0.07, $p<0.01$; β_{item}= -0.70, SE=0.03, $p<0.001$), with an increase in looks to the *cause* location after the IC verb, and marginal by participants but significant by items in the TOP condition (β_{subj}=0.47, SE=0.07, $p=0.07$; β_{item}=0.27, $p<0.005$), with an increase in looks to the *occasion* location after the TOP verb. The pattern of results is in keeping with the claim that verb class induced a shift in participants' expectations with IC verbs yielding a *cause* preference and TOP verbs yielding an *occasion* preference.

As shown in Table 4, for the at-chance participants, the only effect that was reliable across both the participant and items analyses was a main effect of Verb Class, with a higher proportion of looks in the IC condition than the TOP condition. The main effect of Time Window was marginal by participants but significant by items, with more looks after the verb than before. There were trends towards a main effect of Location with more looks to *occasion* than *cause*, which was driven by a trend towards a Location × Verb Class interaction with more looks to *occasion* for TOP verbs and more looks to *cause* for IC verbs, a pattern which was apparent only after the verb, as suggested by the marginal Location × Verb Class × Time Window interaction. The remaining interactions did not reach significance.

In sum, these analyses show that, in keeping with our predictions, the above-chance participants' looking behavior before the verb differed from their looking behavior after the verb: They showed sensitivity to verb class, with more looks to the *cause* location after IC verbs and to the *occasion* location after TOP verbs. The at-chance participants showed only a main effect of Verb Class, with more overall looks in the IC condition than the TOP condition.

To examine the time course of these effects in greater detail, we also conducted a bin-by-bin analysis across the regions before and after the critical coherence-biasing verb. Specifically, during each of a series of temporally defined bins, we examine how the proportion of participants' gazes directed to particular screen locations was conditioned by Verb Class. As an index of how quickly listeners generate expectations about upcoming coherence relations based on the verb in the prompt, we are interested in the time point following the verb when listeners showed a significant influence of Verb Class on their looks to the *cause* and *occasion* locations.

For 100 ms bins beginning 400 ms before the offset of the verb and ending 2800 ms after verb offset, we calculated the proportion of looks to each screen location, and created a series of mixed-effect models (under the same procedures outlined previously) with Location and Verb Class as fixed factors. For each bin, models for data aggregated across participants and across items were computed separately for the above-chance and at-chance participants. Because of the large number of analyses involved, we report the complete set of results in Appendix C. Here, we focus on those analyses that resulted in reliable effects across both participants and items.

Considering the above-chance participants first, there is a recurring main effect of Verb Class with higher proportions of looks in the IC condition than the TOP condition. This can be seen in the -400:-300 ms and -300:-200 ms bins and also continuously from 900 ms to 2100 ms. The early effect may reflect the later IC verb offset compared with the TOP verb offset, thereby providing more time for participants to look *somewhere* in the IC condition; the post-verb-offset effect may reflect a stronger response to IC verbs (favoring the *cause* location) than to the TOP verbs (favoring the *occasion* location more weakly). Crucially, starting at the 900:1000 ms bin, there is a clear interaction between Location and Verb Class that continues for the rest of the post-offset window. Across this time period, above-chance participants looked more to the *cause* location in the IC condition and more to the *occasion* location in the TOP condition.

The at-chance participants show a similar recurring main effect of verb class in bins -400:-300 ms, -300:-200 ms, -100:0 ms, 0:100 ms, 200:300 ms, 600:700 ms, from the 900:1000 ms bin to the 1400:1500 ms bin, from the 1600:1700 ms bin to the 1900:2000 ms bin, and again from the 2200:2300 ms bin to the 2600:2700 ms bin. A Location × Verb Class interaction starts at the 2300:2400 ms bin. Unlike in Experiment 1, the at-chance participants' looking was similar to the above-chance participants', with a *cause* preference in the IC condition and an *occasion* preference in the TOP condition. The difference lies in the time course.

We followed up the Location × Verb Class interaction at each bin to test for the presence of a full crossover interaction. Above-chance participants showed a preference for the *cause* location in the IC condition starting at the 1000:1100 ms bin (β_{subj} = -0.62, SE = 0.07, p < .05; β_{item} = -0.35, SE = 0.03, p < .01); there was no other reliable effect of Location before that bin, and the effect of Location was significant for all subsequent bins. In the TOP condition, the main effect of location favoring the *occasion* location started at the 2000:2100 ms bin (β_{subj} = 0.77, SE = 0.08, p < .05; β_{item} = -0.83, SE = 0.03, p < .001). For the at-chance participants, a main effect of Location for the IC condition favored the *cause* location at the 2700:2800 ms bin (β_{subj} = -0.99, SE = 0.18, p < 0.05; β_{item} = -0.84, SE = 0.12, p < .05), whereas the preference for the *occasion* location in the TOP condition was marginal by participants and significant by items at the 1800:1900 ms bin (β_{subj} = 0.67, SE = 0.09, p = 0.07; β_{item} = 0.87, SE = 0.06, p < .001) and 1900:2000 ms bin (β_{subj} = 0.67, SE = 0.09, p = 0.09; β_{item} = 0.84, SE = 0.06, p < .001).

We also conducted a posthoc analysis to examine participants' performance over the course of the test phase. As noted in the Materials section, all items were presented such that the location of the reappearing ball matched the continuation type that participants heard. Given this consistency, participants could potentially use the items in the test phase to (continue to) master the location~relation mapping. We therefore might see earlier anticipation effects in trials in the second half of the test phase compared to trials in the first half. We therefore conducted bin-by-bin analyses for items from the first half of the test phase separately from those appearing in the second half. Here, we report the earliest bin with a significant Location × Verb Class interaction as an indicator of whether the two halves revealed different time courses of effects. We do not list the statistics for subsequent bins for reasons of space, though it should be noted that certain analyses revealed some marginal or non-significant bins after the first significant interaction, particularly for at-chance participants. For the above-chance participants, trials in the first half of the test phase first show a reliable Location × Verb Class interaction in the 1200:1300 ms bin ($\beta_{\text{subj}} = 0.87$, $SE = 0.12$, $p < 0.05$; $\beta_{\text{item}} = 0.75$, $SE = 0.11$, $p < 0.01$), whereas trials in the second half show a reliable interaction in the 400:500 ms bin ($\beta_{\text{subj}} = 0.60$, $SE = 0.10$, $p < 0.05$; $\beta_{\text{item}} = 1.06$, $SE = 0.15$, $p < 0.05$). For at-chance participants, trials in the first half show a significant Location × Verb interaction in only two bins, 1800:1900 ms ($\beta_{\text{subj}} = 1.17$, $SE = 0.16$, $p < 0.05$; $\beta_{\text{item}} = 1.14$, $SE = 0.14$, $p < 0.001$) and 2200:2300 ms ($\beta_{\text{subj}} = 0.86$, $SE = 0.17$, $p < 0.001$; $\beta_{\text{item}} = 1.0$, $SE = 0.18$, $p < 0.05$). Later trials showed no reliable interaction for at-chance participants.

These analyses indicate that, at least for the above-chance participants, the critical interaction shifted 800 ms earlier from the first half of the test trials to the second half, suggesting that those participants became more accustomed to the task over the course of the experiment or more familiar with the location~relation mapping. Either way, the above-chance participants' anticipatory looking in the second half frequently started before the prompt was complete. For the at-chance participants, there is no evidence that their understanding of the relevant mapping grew stronger from the beginning to the end of the test trials.

5.3 Discussion

The results from Experiment 2 further demonstrate how this learning-and-looking paradigm can be used to probe listeners' online coherence biases. Specifically, the results show that verb class is one type of cue that drives coherence expectations, in line with previous work using offline methodologies (Kehler et al., 2008). Participants who were above chance on the post-training assessment showed the predicted pattern during the test phase whereby a *cause*-biasing cue (an IC verb) yielded anticipatory looks to the *cause* location and an *occasion*-biasing cue (a TOP verb) yielded anticipatory looks to the *occasion* location. This Location × Verb

Class interaction emerged only after the offset of the verb, showing that the coherence-biasing cue had an impact on participants' looking behavior. Bin-by-bin analyses tested the time course of the effect, showing anticipation within the first second following the coherence-biasing cue for above-chance participants and later effects for at-chance participants. The effects for the at-chance participants were delayed by roughly a second, and hence the analysis showed no reliable result for that group. However, in the bin-by-bin analysis, the direction of the effects echoes the above-chance participants. Both groups showed evidence of having learned the location~relation mapping and of being sensitive to verb class.

Recall that in Experiment 1, participants' looking behavior was difficult to interpret because the coherence-biasing cue was itself a coherence marker (a connective). The coherence-biasing cue used in Experiment 2 avoids this difficulty by relying on a set of verbs whose lexical semantic meanings impose no explicit constraints on coherence relations; it is only the coherence biases that participants associate with these verbs that set up an expectation for a particular type of continuation in the ensuing discourse. The pattern of looking behavior supports the conclusion that a local cue in one sentence can influence listeners' expectations about the coherence relation that will be established with an upcoming sentence.

The results also help clarify the looking behavior observed during the first proposition of the prompts in Experiment 1. Since many of those clauses contained IC verbs, the observed bias to the *cause* location likely reflects a verb-driven response. In this light, the early looking behavior in Experiment 1 can be taken as anticipatory, which suggests that even in time windows in which the task itself does not encourage anticipation, a natural reflex of participants' language processing includes a categorization of coherence-biasing cues and an anticipation regarding the role that upcoming material will play in establishing discourse coherence.

Regarding the strength of the verb-driven biases, the results revealed stronger looking preferences in the IC condition than in the TOP condition. This can likely be attributed to the strength of the respective *cause* and *occasion* biases of the verbs themselves. Using evidence from story continuations, Kehler et al. (2008) reported that IC context sentences strongly favor an upcoming Explanation continuation (60% of continuations), whereas TOP context sentences yield more variation in continuation type, with the most common type being Occasion (44% of continuations). In our results, it appears that the strength of anticipation reflects the strength of the coherence-biasing cue, suggesting a fine-grained sensitivity to the probabilistic nature of coherence-biasing cues.

Lastly, the time course of the effects shows that anticipation happens quickly, often before the end of the sentence. The post-hoc analysis of trials from the first and second halves of the test phase suggests that participants continued to learn about the location~relation mapping over the course of the experiment, allowing their looking behavior to speed up. In the second half of the experiment, above-chance participants showed the predicted pattern of anticipatory looking as early as 400-500 ms after verb offset, well before the end of the sentence.

6. General Discussion

The identification of intersentential coherence relations is one of the key inferences that comprehenders must make in interpreting a discourse. The results from Experiment 1 show that, for listeners who learn the relevant location~relation mapping, their looking behavior is linked to the presence of a coherence-biasing cue, namely an intersentential connective (“because” or “so”). The results from Experiment 2 demonstrate that the establishment of coherence relations is expectation-driven, with comprehenders using an available cue from one sentence to anticipate the relation that is likely to hold between the current sentence and the next. Previous studies using offline methodologies like sentence completion have shown that comprehenders have strong biases about the kinds of coherence relations that are likely to hold given biasing material in the preceding context (Kaiser, 2009; Kehler et al., 2008; Rohde, et al. 2007; Simner & Pickering, 2005). However, it is not clear from such work when and how strongly such expectations are generated in the course of comprehension. Here we report online evidence to suggest that local linguistic cues can indeed guide high-level expectations about intersentential coherence relations.

We were able to demonstrate these early expectations through a novel eyetracking paradigm modeled on McMurray and Aslin's (2004) infant category-learning studies. After undergoing a training phase that associated different continuation types with specific locations on a visually presented "classifier," participants' looks reflected coherence-biasing cues in the prompts. Importantly, these patterns of looking behavior occurred before participants heard the continuation sentences or saw any visual cue regarding the type of continuation to come.

6.1 Strength and speed of response

Our eyetracking paradigm also allowed us to examine the time course of these expectations: For participants who successfully learned the location~relation mappings during training, their looking pattern shifted soon after the offset of the coherence-biasing cue. In Experiment 1, it was unclear whether the pattern of fixations reflected the integration of the connective itself or, rather, coherence anticipation in line with Simner & Pickering's Satisfied Gap Hypothesis. In Experiment 2, however, the materials were more clearly designed to test anticipation, and we saw evidence of anticipatory looking 900-1000 ms post verb for successful learners. Our analysis of the data from the second half of the test phase shows evidence of anticipatory looking to verb-congruent regions as early as 400-500 ms post verb. These results suggest that inferencing about likely coherence relations between pairs of sentences is an online process, starting before one sentence is complete and well before both sentences are available.

Such an online expectation-driven account stands in contrast to models such as Clausal Integration (Garnham et al., 1996; Stewart et al., 2000), which posit that the identification of discourse relations requires the availability of both relevant sentences in their entirety. Evidence supporting Clausal Integration comes from work on anaphor resolution and typically involves manipulations with IC verbs. Using measures such as visual probe recognition (Garnham et al., 1996) and reading time (Stewart et al., 2000), work supporting late integration reports IC effects at the end of sentence processing but few effects on early pronoun resolution. The methods and materials used in these studies have been criticized, however (e.g., Koornneef & Van Berkum, 2006; Cozijn, Commandeur, Vonk, & Noordman, 2011), for failing to provide conditions that allow an adequate test of early "focusing" accounts of implicit causality in pronoun resolution. In subsequent work addressing these concerns, evidence from eye fixations on critical sentence regions during self-paced reading (Koornneef & Van Berkum, 2006), ERP responses to incongruent pronouns (Van Berkum, Koornneef, Otten, & Nieuwland, 2007), and eye movements toward referents of ambiguous pronouns in visual world paradigms (Cozijn et al., 2011; Pyykkönen & Järvikivi, 2010) have obtained evidence consistent with early IC effects. Our current results extend these findings beyond the identification of pronoun antecedents to the inference of coherence relations. Additionally, given that prior models (including recent studies on early focusing) have measured the time course of pragmatic interpretation by exclusively targeting IC contexts, one might conclude that comprehenders only resort to online inferencing when faced with the strong biases associated with IC verbs. Our use of other non-IC cues points to the pervasiveness of such inferencing during online processing.

As a testament to the strength of coherence expectations and the speed with which they are generated during discourse comprehension, even participants in Experiment 2 who were at chance on the training assessment showed evidence of anticipating verb-congruent coherence relations. Even so, it appears as if the at-chance participants did need to hear most, if not all, of the prompt before reliably looking toward a particular screen location in anticipation of the next sentence. For those participants, generating the relevant expectations may have required more time or input or they may have been slower to convert their coherence expectation from a relation category into a spatial location, given their unreliably learned location~relation mapping.

Although the participants' behavior indicates that they learned more about the mapping over the course of the experiment, in the end very few participants across either experiment were able to articulate that mapping. We take this to suggest that the coherence relations we used and the types of pragmatic reasoning required to infer such relations were largely implicit. Comprehending a discourse as more than a sequence of arbitrary sentences depends on the important assumption that sentences appear together because they relate in meaningful ways. The inferences required to identify those relations and infer a coherent discourse need not be expressed overtly. As we have shown, comprehenders need not even be able to articulate the inferred relations in order to anticipate their appearance in subsequent discourse.

6.2 Inference of coherence categories

Not only did participants vary in their ability to articulate the classifier categories, but those who could articulate a pattern also varied in the categories they inferred. This raises the question of what features participants were responding to in their (usually implicit) categorization of the prompt and continuation sentences: Many of the training and test items are compatible with a categorization scheme based on either causal reasoning or temporal progression. Research on the nature of discourse coherence has long acknowledged that causal and temporal interpretations are deeply intertwined (Lascarides & Asher, 1993). Descriptions of causes are likely to describe events that happened in the past, whereas descriptions of consequences move forward in time. In other words, the *cause*-vs-*consequence/occasion* categories are synonymous in many contexts with backward/forward movement in time. The interpretation of temporal progression may depend on the inference of causation, although the reverse is not true: Causation can be inferred irrespective of temporal marking in certain cases. For example, the passage "*John likes Mary. She is beautiful.*" requires causal reasoning to infer that the stative in the second sentence is an explanation of the stative in the first, but those two sentences do not describe states of affairs that are themselves temporally ordered (since they do not describe temporally constrained events at all).

The majority of participants who could articulate a pattern did so with reference to causality, but there were some who mentioned time instead. Although Experiment 1 targeted the *cause*/*consequence* categories from Simner and Pickering (2005), Experiment 2 targeted a *cause*/*occasion* distinction. Our norming study of the *consequence* and *occasion* training items confirmed that the Experiment 2 training did indeed provide a weaker basis than the Experiment 1 training for the inference of causal categories (compared to a purely temporal categorization). That said, participants in the norming study inferred causality more often than not, in keeping with claims that causal connections in text are often preferred and facilitate comprehension and recall (Trabasso & Sperry, 1985; Trabasso & van den Broek, 1985). As noted in both experiments, the materials used in the training and test phases included items in which time and causality were conflated, while other items described states of affairs or habitual events that resist temporal ordering. This latter category of items may have encouraged participants to articulate the location categories in terms of causality. Under either the causal or temporal interpretation of the categories, the results demonstrate that comprehenders use local cues to anticipate a relationship between the current sentence and an upcoming sentence.

6.3 Implications for models of sentence processing

The relevance of our results to existing models of sentence processing follows from the established effects of discourse-level biases on a variety of well-studied psycholinguistic phenomena. These phenomena range from sentence-internal syntactic processing (Altmann & Steedman, 1988; Crain & Steedman, 1985; Ni, Crain, & Shankweiler, 1996; Rohde et al., 2011; Sedivy 2002; Van Berkum et al., 1999) to between-sentence ellipsis interpretation (Kehler, 2000; Kehler, 2002) and pronoun resolution (Kaiser, 2009; Kehler, 2002; Kehler et al., 2008; Rohde, Kehler, & Elman, 2007; Wolf, Gibson, & Desmet, 2004). Within the sentence, studies in syntactic processing have primarily focused on the role of an upcoming clause to restrict reference as a factor in the resolution of local syntactic ambiguity. As Crain and Steedman (1985) and others have shown, the syntactic attachment point of linguistic material following a noun phrase is modulated by the status of the noun phrases' referent—if the referent is not uniquely identifiable in the discourse context, upcoming material is likely to be interpreted as modifying that noun phrase (as opposed to attaching to some other position in the syntactic structure) precisely because the upcoming material may serve to restrict reference, particularly in contexts with definite determiners that impose a requirement of referential uniqueness. The more recent results of Rohde et al. (2011) show that comprehenders' expectations about upcoming post-nominal material is not limited to reference restriction, but may also reflect expectations for an explanation of an eventuality described earlier in the sentence. As such, these previous studies have established online evidence for comprehenders' sensitivity to discourse-level information and even, in the case of Rohde et al., online evidence for comprehenders' anticipation of upcoming explanation coherence relations. However, in contrast to this previous work focusing on the repercussions of discourse-level expectations for other sentence-processing phenomena, our study directly assesses the generation of expectations about coherence relations.

As with sentence-internal processing, existing work on between-sentence phenomena like verb-phrase ellipsis and pronoun resolution has similarly identified a role for the operative coherence relation in the interpretation of elided and ambiguous material. This previous work, however, has not addressed online coherence expectations, focusing instead on offline methodologies or contexts with overt intersentential connectives. A better understanding of coherence establishment, and the cues that guide it, will hopefully inform work on these types of between-sentence phenomena. Prior reliance on phenomenon-specific studies moreover fails to capture the fact that coherence is relevant to a variety of phenomena and is, we believe, integral to language processing as a whole, not just to the resolution of ambiguous pronouns or ambiguous structures. With the method presented here, it is possible to probe coherence expectations more directly, rather than through the lens of other coherence-sensitive phenomena. Lastly, our results bring the pragmatic process of coherence establishment in line with a variety of other linguistic phenomena that show early, expectation-driven effects.

7. Conclusion

The results presented here demonstrate that listeners' anticipatory looks during sentence processing can be used to test their pragmatic expectations about discourse coherence relations. Using a novel eyetracking paradigm, we trained participants to associate particular spatial locations with particular coherence relations. Subsequently, listeners' eye movements during sentences containing coherence-biasing cues revealed preferential looking to the locations that had been associated with *cause* and *consequence/occasion* continuations. Experiment 1 showed, for above-chance participants, differential looking following coherence-signaling connectives, as well as preferential looking to the *cause* location during initial clauses containing implicit causality verbs. Experiment 2 showed anticipatory looking following verbs with different coherence biases. Importantly, all of these patterns emerged prior to subsequent continuations of the discourse. These results extend existing research showing expectation-driven processing at other levels of linguistic processing. Whereas that previous work demonstrated that comprehenders generate expectations about upcoming sounds, words, and syntactic structures within the sentence, the findings reported here point to expectations about relationships that hold between entire sentences. Moreover, our novel methodology provides new opportunities for examining the types of cues that comprehenders use to establish intersentential relationships and the types of discourse structures they build as they infer coherent meaning from a series of juxtaposed sentences.

Acknowledgements

This research was supported by an Andrew W. Mellon postdoctoral fellowship to the first author. The results from this paper have been presented at the 23rd CUNY Conference on Human Sentence Processing, New York, NY, March 2010. We thank Sudha Arunachalam, Brady Clark, Matt Goldrick, Chigusa Kurumada, Elizabeth Mazzocco, Daniel Slaten, Michael Tanenhaus, and Masaya Yoshida for helpful discussion. We also thank research assistant Ronen Bay. We are grateful for the comments and suggestions from three anonymous reviewers.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, *71*, 247-264.

Altmann, G. T. M, & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.

Arnold, J. E. (2001). The effects of thematic roles on pronoun use and frequency of reference. *Discourse Processes*, *31*, 137-162.

Asher, N. (1993). *Reference to Abstract Objects in Discourse*. *Studies in Linguistics and Philosophy*. Dordrecht: Kluwer Academic Publishers.

Asher, N., & Lascarides, A. (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.

Au, T. K. (1986). A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. *Journal of Memory and Language*, *25*, 104-122.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457-474.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Bates, D., Maechler, M., & Bolker, B. (2011). *lme4: Linear mixed-effects models using S4 classes*.

Brown, R. & Fish, D. (1983). The psychological causality implicit in language. *Cognition*, *14*, 237-273.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In D. Dowty, L. Karttunen, and A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 443–467). Cambridge.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language, 47*, 30-49.

Cozijn, R., Commandeur, E., Vonk, W., & Noordman, L. G. M. (2011). The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language, 64*, 381-403.

Delong, K. A., Urbach, T. A., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*, 1117-1121.

Desmet, T., De Baecke, C., & Brysbaert, M. (2002). The influence of referential discourse context on modifier attachment in Dutch. *Memory and Cognition*, *30*, 150–157.

Ferretti, T., Rohde, H., Kehler, A., & Crutchley, M. (2009) Verb aspect, event structure, and coreferential processing. *Journal of Memory and Language, 61*, 191-205.

Ferstl, E. C., Garnham, A. & Manouilidou, C. (2011). Implicit causality bias in English: A corpus of 300 verbs. *Behavior Research Methods, 43*, 124-135.

Filoteo, J. V., Maddox, W. T., Simmons, A. N., Ing, A. D., Cagigas, X. E., Matthews, S., & Paulus, M. P. (2005). Cortical and subcortical brain regions involved in rule-based category learning. *NeuroReport*, *16*, 111-115.

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, 31, 931.

Garnham, A., Traxler, M., Oakhill, J., & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language*, *35*, 517–543.

Garvey, C., & Caramazza, A. (1974). Implicit causality in verbs. *Linguistic Inquiry*, *5*, 459-464.

Grimes, J. (1975). *The thread of discourse*. The Hague: Mouton.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hobbs, J. R. (1979). Coherence and coreference. *Cognitive Science, 3*, 67–90.

Hobbs, J. R. (1990). *Literature and cognition*. Stanford, CA: CSLI. Lecture Notes 21.

Kaiser, E. (2009). Effects of anaphoric dependencies and semantic representations on pronoun interpretation. In S.L. Devi, A. Branco and R. Mitkov (Eds.), *Anaphora Processing and Applications (Selected papers from the 7th Discourse Anaphora and Anaphor Resolution Colloquium)*. pp.121-130. Heidelberg: Springer.

Kaiser, E. (2012). Taking action: a cross-modal investigation of discourse-level representations. *Frontiers in Psychology*, 3:156.

Kaiser, E., & Trueswell, J.C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, *94*, 113-147.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*, 33-156.

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, *32*, 37-55.

Kehler, A. (2000). Coherence and the resolution of ellipsis. *Linguistics and Philosophy*, *23*, 533-575.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25*, 1–44.

Kehler, A., & Rohde, H. (2014). A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics*.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition*, *95*, 95-127.

Koornneef, A. W., & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye-tracking. *Journal of Memory and Language*, *54*, 445–465.

Lascarides, A., & Asher, N. (1993) Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy, 16*, 437-493.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126-1177.

Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, *8*, 243-281.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. London: Chapman and Hall.

McKoon, G., Greene, S., & Ratcliff, R. (1993). Discourse models, pronoun resolution, and the implicit causality of verbs. *Journal of Experimental Psychology*, *19*, 1040-1052.

McMurray, B., & Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, 6, 203-229.

Most, R. B., & Saltz, E. (1979). Information structure in sentences: New information. *Language and Speech*, *22*, 89-95.

Ni, W., Crain, S., & Shankweiler, D. (1996). Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes*, *11*, 283–334.

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A. & Webber, B. (2008) The Penn Discourse Treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation* (LREC). Marrakech, Morocco.

Pyykkönen, P., & Järvikivi, J. (2010). Activation and persistence of implicit causality information in spoken language comprehension. *Experimental Psychology*, *57*, 5-16.

Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics*, 49: Papers in Semantics.

Rohde, H. & Kehler, A. (2008). The bidirectional influence between coherence establishment and pronoun interpretation. Poster presented at the 21st Annual CUNY Conference on Human Sentence Processing. Chapel Hill.

Rohde, H., & Kehler, A. (2009). The interaction of structural and semantic biases in coherence and coreference. Poster presented at the 22nd Annual CUNY Conference on Human Sentence Processing. Davis, CA, March 26-28.

Rohde, H., Kehler, A., & Elman, J. L. (2006). Event structure and discourse coherence biases in pronoun interpretation. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 697-702. Vancouver: Cognitive Science Society.

Rohde, H., Kehler, A., & Elman, J. L. (2007). Pronoun interpretation as a side effect of discourse coherence. In *Proceedings of the 29th Annual Conference of the Cognitive Science Society,* 617-622. Nashville, TN: Cognitive Science Society.

Rohde, H., Levy, R., & Kehler, A. (2011). Anticipating explanations in relative clause processing. *Cognition*, *118*, 339-358.

Sedivy, J. C. (2002). Invoking discourse-based contrast sets and resolving syntactic ambiguities. *Journal of Memory and Language, 46*, 341–370.

Simner, J., & Pickering, M. J. (2005). Planning causes and consequences in discourse. *Journal of Memory and Language, 52*, 226-239.

Stevenson, R., Crawley, R., & Kleinman, D. (1994). Thematic roles, focusing and the representation of events. *Language and Cognitive Processes, 9*, 519-548.

Stewart, A. J., Pickering, M. J., & Sanford, A. J. (2000). The time course of the influence of implicit causality information: Focus versus integration accounts. *Journal of Memory and Language*, *42*, 423–443.

Tanenhaus, M.K., Spivey, M.J., Eberhard, K.M., & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632-1634.

Tanenhaus, M. K., & Trueswell, J. C. (2005). Eye movements as a tool for bridging the language-as-product and language-as-action traditions. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-action and language-as-product traditions* (pp. 3-37). Cambridge, MA: MIT Press.

Trabasso, T., & Sperry, L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, *24*, 595-611.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*, 612-630.

Traxler, M., Sanford, A. J., Aked, J., & Moxey, L. (1997). Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 88-101

Tyler, J. (2012). Discourse Prosody in Production and Perception. PhD Dissertation, University of Michigan.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*, 147–182.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*, 443–467.

Van Berkum, J. J. A., Koornneef, A. W., Otten, M. & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research, 1146*, 158-171.

Weiler, J. A., Bellebaum, C., & Daum, I. (2008). Aging affects acquisition and reversal of reward-based associative learning. *Learning & Memory*, *15*, 190-197.

Wolf, F., Gibson, E., & Desmet, T. (2004). Discourse coherence and pronoun interpretation. *Language and Cognitive Processes*, *19*, 665-75.

Appendix A

Experiment 1 items: The *consequence*-biasing prompt (containing "so") is listed first, followed by the *cause*-biasing prompt (containing "because"). The prompts are followed first by the *cause* continuation and then by the *consequence* continuation. Participants heard only one prompt and one continuation for each item.

Beryl admired John so she applauded him.
Beryl delighted John because she applauded him.
    His incredible piano playing had always impressed her.
    He grinned proudly at the applause.
Thomas disliked Debbie so he criticised her.
Thomas provoked Debbie because he criticised her.
    Her loud voice had recently been getting on his nerves.
    She retaliated by insulting his mother.
Arnold was worried for Jennifer so he helped her.
Arnold was thanked by Jennifer because he helped her.
    She had been failing her math class as well as English and Chemistry.
    She gave him a huge hug out of gratitude.
Jonathan was scared of Nicola so he obeyed her.
Jonathan pleased Nicola because he obeyed her.
    She commanded respect and was impossible to ignore.
    She became more and more bossy.
Ian suspected Amy so he accused her.
Ian embarrassed Amy because he accused her.
    She had been the only one home when the fire started.
    She denied the accusation.
Beth liked Rob so she bantered with him.
Beth impressed Rob because she bantered with him.
    He had always appreciated people who were willing to disagree with him.
    He later teased her for her flirtatious bantering.
Carol wanted Trevor so she chased him.
Carol stressed Trevor out because she chased him. (changed from "caught")
    He was playing hard-to-get.
    He told her to stop pestering him.
Claire trusted Gordon so she confided in him.
Claire surprised Gordon because she confided in him.
    He was known for guarding people's secrets well.
    He kept her secret til the day he died.
Emma noticed Adam arriving so she greeted him.
Emma made a good impression on Adam because she greeted him.
    He had arrived late and was new to the school.
    He smiled warmly at her.

Jane detested Ted so she harassed him.
Jane sickened Ted because she harassed him.
    He was terribly timid and spineless.
    He responded in kind and attacked her.
Ted hated Jane so he harmed her.
Ted frightened Jane because he harmed her.
    She had made the mistake of bringing up his drinking habit.
    She reported the incident to the police.
Jennifer was angry with Arnold so she hit him.
Jennifer was disliked by Arnold because she hit him.
    He had been mocking her in front of everyone.
    He nursed a grudge for weeks.
Joan hated Barry so she hurt him.
Joan frightened Barry because she hurt him.
    She had repressed her anger for years until it exploded.
    He stared in disbelief at the woman he thought he loved.
Kate disliked Donald so she insulted him.
Kate offended Donald because she insulted him.
    She couldn't take his arrogant attitude any more.
    He walked away quietly with her insults still ringing in his ears.
Donald had enough of Kate so he interrupted her.
Donald offended Kate because he interrupted her.
    She always wants to be the center of attention.
    She stopped mid-sentence with a surprised look on her face.
Stephen wanted to make Liz relax so he joked with her.
Stephen relaxed Liz because he joked with her.
    She had been tense all day.
    She laughed despite her bad mood.
Lorna was furious with Henry so she kicked him.
Lorna was in trouble with Henry because she kicked him.
    She was throwing a temper tantrum.
    He kicked her back.
Peter got angry with Rose so he pushed her.
Peter scared Rose because he pushed her.
    He had started losing his temper more and more often.
    She fell backwards against the wall.
Sue hated Luke so she slandered him.
Sue infuriated Luke because she slandered him.
    She wanted revenge after their breakup.
    He took her to court.
Luke disliked Sue so he snubbed her.
Luke angered Sue because he snubbed her.
    She is nosy and hates to be ignored.
    She kept calling and leaving messages for him for weeks afterwards.
Jack liked Susan so he stared at her.
Jack annoyed Susan because he stared at her.

He was hopelessly in love with her.
She started to feel uncomfortable.
Susan needed Jack so she stopped by to see him.
Susan startled Jack because she stopped by to see him.
She wanted to get his advice.
He agreed to help her just this once.
Suzanne admired Richard so she supported him.
Suzanne pleased Richard because she supported him.
He was a starving artist and needed the money.
He felt guilty for taking her money.
Veronica was worried for Callum so she warned him.
Veronica saved Callum because she warned him.
He was going to be fired under false pretenses.
He ignored her warnings and insisted everything would be okay.

Appendix B

Experiment 2 items: Each prompt is presented followed first by the *cause* continuation and then by the *occasion* continuation. Participants heard only one continuation for each item.

Implicit Causality Items

Dawn amazed Malcom.
    She was playing the piano with her eyes closed.
    He applauded her talents.
Burt amused Gwen.
    He told her a joke about elephants.
    She laughed out loud.
Nancy annoyed Doug.
    She talked on the phone all day.
    He stomped out of the room.
Kevin apologized to Bethany.
    He had lied to her about his past.
    She accepted his apology.
Pamela bored Felix.
    She told endless stories about her pet poodle.
    He fell asleep while she was talking.
Chad charmed Monica.
    He told her riveting stories about life in New York.
    She agreed to go out with him on a date.
Caitlin offended Alan.
    She made fun of his mother.
    He glared back at her.
Kyle confessed to Joyce.
    He felt guilty about lying to her.
    She forgave him for lying to her.
Valerie deceived Roy.
    She didn't want him to know her real name.
    He was hurt by her lies when he found out.
Albert disappointed Eliza.
    He was late for their date yet again.
    She refused to speak to him.
Throughout the evening, Betsy aggravated Henry.
    She talked nonstop.
    He left the party early to avoid her.
During high school, Austin exasperated Dorothy.
    He adamantly refused to take steps to improve his grades.
    She decided to hire a tutor to help him study better.
At the cocktail party, Alexandra fascinated Simon.
    She had grown up in Australia.

He asked for her phone number.
Over the years, Greg frightened Barbara.
He had a terrible temper.
She finally reported him to the police.
In the cafeteria, Megan humiliated Grant.
She showed everyone the poem he wrote.
He blushed with embarrassment.
James infuriated Sandra on their second date.
He argued politics whenever he got the chance.
She decided to break up with him.
Genevieve inspired Ronald during college.
She worked hard and had received a scholarship.
He decided to pursue a career in journalism.
Ray intimidated Tiffany at the bowling alley.
He was bowling strike after strike.
She became so nervous that she could barely pick up her ball.
Darleen scared Conor one night.
She liked to drive at night with no headlights.
He warned her that her behavior was a risk to her health.
Logan surprised Ashley one morning.
He had purchased roundtrip tickets to London.
Her jaw nearly hit the floor.
Abe assisted Eileen.
She was having trouble with her homework.
She thanked him for his assistance.
Lucy blamed Bruce.
He was the one who left the doors unlocked.
He denied being responsible.
Mary comforted Joseph.
He was scared of the dark.
He finally was able to fall asleep despite his nightmare.
Theo congratulated Miriam.
She had just won first prize in the gymnastics competition.
She grinned and walked up to receive the first prize ribbon.
Nick corrected Courtney.
She had misspelled the state capital.
She quickly fixed the errors.
Stacie detested Hank.
He thought women shouldn't be allowed to vote.
She asked everyone to stop talking to him.
Eva envied Calvin.
He drove a brand new Porsche.
She asked him how he got so rich.
Rob hated Natalie.
She never washed her socks.
He told her to get out of his face.

Todd helped Brooke.
    She was having trouble reading the map.
    She wrote him a thank you note.
Scott mocked Ruth.
    She wearing her shoes on the wrong feet.
    He then told her to take off the silly hat with elephant ears.
In recent weeks, Rebecca feared Oliver.
    He was blackmailing her about her past life.
    She requested a restraining order.
On Michigan Avenue, Martin noticed Celeste.
    She was standing across the street waving.
    He waved hello.
After the first day of school, Ebony pacified Neal.
    He was panicking about the amount of work to do.
    She then offered to take him out for ice cream.
Standing on the stage, Brandon praised Kate.
    She had won the state violin contest for the third year in a row.
    He proceeded to appoint her as the chairwoman.
Vanessa reproached Walter in the kitchen.
    He was eating dessert before dinner.
    He apologized for sneaking a cookie from the cookie jar.
Arthur scolded Patricia in the hallway.
    She had put thumbtacks on the teacher's chair.
    He then sent her to the principal's office.
Emily stared at Chuck at the swim meet.
    He was the buffest guy she had ever seen.
    She walked over and boldly introduced herself.
Luis thanked Debbie on Christmas Eve.
    She had washed all the dishes after the big family dinner.
    He promised to do his share next time.
Stephanie trusted Curtis for years.
    He had been there for her during her illness.
    She called him to ask advice about grad school.
Jasper valued Leah during the campaign.
    She consistently provided good advice without judgment.
    He invited her to join his team.

Transfer of Possession Items

Caleb threw Claire a hat.
    He was worried that she was cold.
    They ran out the door into the snow.
Elizabeth delivered a letter to Darrell.
    She was the local mail carrier.
    He opened it with curiosity.

Ian slugged a line drive to Kimberly.
He hoped to make it to first base.
He sprinted towards first base.
Candice wheeled a large bicycle to Richard.
They had planned to go biking as their first date.
He climbed on and rode off.
Bill slapped a beachball to Michelle.
He wanted her to have fun at the beach.
She dove for the ball and fell into a wave.
Carol smacked a ball to Dylan.
She was angry at him for not paying attention to the game.
He stumbled awkwardly as he reached for it.
Luke spun Helen a top.
He was bored sitting in the playroom with his sister.
She giggled at the spinning colors.
Alex socked a ball to Ed.
She was trying to get the runner out on first.
He called 'foul'.
Heidi shipped Eric a package.
She thought he would like some cookies from home.
He wrote a thank you note.
Jeremy served Eve chili.
He wanted her to try his famous recipe.
She gobbled it up.
Meredith shot a puck to Xavier.
They were on the same team and only together could make the goal.
He skated quickly across the rink.
Graham shoved a mass of papers to Linda.
He needed help filing all the paperwork.
She sighed at the huge amount of work.
Lisa brought Tony a cup of tea.
He had been complaining of a sore throat.
She said 'good morning' as she set the tea down.
Bruno dropped a roof shingle to April.
They were putting down a new roof before winter.
She warned 'be careful!'
Marilyn rolled a toy truck to Jack.
She wanted someone to play with.
He grabbed another toy out of the toy chest.
Hal transmitted a message to Emma.
She was waiting to be told where to go.
He checked his watch to confirm the time.
Danielle pitched a ball to William.
They were practicing for the big game.
She listened to the roar from the stands.
Max fed Veronica a piece of chocolate.

He hoped to win her over with chocolate.
She smiled lovingly at him.
Christina flipped a queen of spades to Victor.
He was playing poker and had asked for another card.
She declared the game over.
Andrew propelled a beer can to Cindy.
He was hoping to impress her with his aim.
She yelled 'watch out!'
Zoe towed an old car to David.
She wanted him to fix the muffler.
He made fun of the paint job.
Justin kicked a soccer ball to Alissa.
He hoped she would not let it go out of bounds.
She trapped the ball with her foot.
Kristen pushed a shopping cart to Nigel.
He was holding several bags that were about to drop.
He placed some arugula in the cart.
Charles skipped a beach ball to Jill.
He wanted to make her laugh.
She laughed as the ball fell in the water.
Margaret whacked a wiffleball to Brad.
They were playing ball to pass the time.
He sighed at her awful hit.
Arnold batted a grounder to Edith.
He knew she was hopeless at catching grounders.
He ran towards first base.
Amy passed Roger a sandwich.
She wanted him to try liverwurst.
He got up to get a soda.
Vincent swatted a tennis ball to Grace.
He was testing out his new racket.
She watched it sail over her head.
Kara heaved a box to Heath.
She wanted him to put the box in the attic.
He ripped the top open to see what was inside.
Brian hauled a wheelbarrow to Denise.
They were hoping to replant the garden that Sunday.
She opened the barn door.
Angela chucked Dustin a measuring tape.
She wanted him to help measure the table.
He measured the length of the table.
Barry hit an easy fly ball to Nicole.
He was teaching her to play softball.
She caught it despite having the sun in her eyes.
Erika nudged a microphone to Drew.
She wanted him to make a speech.

He took it and stared out at the audience.

Marvin hurled a brick to Janice.

They were having a fight.

She managed to duck just in time.

Andrea snapped a Frisbee to Joel.

She was trying to amuse her brother in the backyard.

He then dashed across the field.

Blane lifted a box to Ethel.

They were putting Christmas decorations back into the attic.

She loaded the box into the moving van.

Heather bounced a basketball to Josh.

He was the only open player on the court.

She noticed that the ball was deflated.

Jeff tossed Rosalind a dishtowel.

He wanted help doing the dishes.

She grabbed it just as the oven timer went off.

Amanda lobbed a football to Anthony.

They were trying to keep the ball away from their father.

He fell just as he reached the 30-yard line.

Mitchel handed Hannah a timecard.

She was new on the job and didn't know she needed to punch in.

She asked about the upcoming union meeting.

Appendix C

Mixed model coefficient estimates and standard errors for Experiment 2, computed separately by performance on the post-training assessment and analysis bin (100 ms bins; verb offset = 0 ms)

		Above-Chance (N = 26)				At-Chance (N = 16)			
		Participants		Items		Participants		Items	
Bin (ms)	Fixed Effect	β	*SE*	β	*SE*	β	*SE*	β	*SE*
-400:-300	Location	0.02	0.07	0.23*	0.03	-0.25	0.11	0.02	0.1
	Verb Class	-0.48**	0.03	-0.34**	0.03	-0.95**	0.12	-0.46**	0.09
	Location × Verb Class	-0.2	0.08	0.2	0.07	0.46	0.16	-0.04	0.2
-300:-200	Location	0.11	0.12	0.3**	0.03	-0.27	0.11	-0.08	0.09
	Verb Class	-0.22*	0.03	-0.16*	0.03	-0.78**	0.1	-0.37*	0.08
	Location × Verb Class	-0.35	0.08	0.01	0.07	-0.69	0.17	-0.33	0.19
-200:-100	Location	0.07	0.12	0.27**	0.03	-0.19	0.12	0.02	0.07
	Verb Class	-0.22	0.03	-0.11	0.03	-0.49*	0.07	-0.23†	0.07
	Location × Verb Class	-0.32	0.07	0.02	0.06	-0.2	0.13	0.12	0.15
-100:0	Location	0.04	0.11	0.23**	0.03	-0.01	0.12	0.27†	0.08
	Verb Class	-0.14	0.04	-0.05	0.02	-0.48*	0.06	-0.28*	0.06
	Location × Verb Class	-0.3	0.08	0.03	0.05	0.03	0.11	0.27	0.15
0:100	Location	0.03	0.11	0.14†	0.03	-0.05	0.11	0.35*	0.07
	Verb Class	-0.24*	0.03	-0.15†	0.03	-0.42**	0.05	-0.25*	0.06
	Location × Verb Class	-0.16	0.08	-0.1	0.05	0.1	0.09	0.23	0.13
100:200	Location	-0.01	0.11	0.03	0.03	-0.06	0.1	0.41**	0.07
	Verb Class	-0.13	0.03	-0.12	0.02	-0.16	0.04	-0.23†	0.06

	Location × Verb Class	-0.06	0.07	-0.13	0.06	0.62^{*}	0.09	0.17	0.13
200:300	Location	-0.18	0.1	0	0.03	0.09	0.08	0.32^{*}	0.06
	Verb Class	-0.08	0.03	-0.15^{*}	0.02	-0.37^{*}	0.05	-0.3^{*}	0.06
	Location × Verb Class	0.28	0.07	0.08	0.05	$0.6^{\dagger}$	0.09	0.29	0.12
300:400	Location	-0.17	0.09	-0.05	0.02	0.15	0.1	0.35^{**}	0.06
	Verb Class	-0.11	0.02	-0.08	0.02	-0.4^{**}	0.05	$-0.19^{\dagger}$	0.05
	Location × Verb Class	0.22	0.06	0.14	0.05	0.29	0.08	0.1	0.11
400:500	Location	-0.19	0.08	-0.02	0.03	0.06	0.1	0.26^{*}	0.05
	Verb Class	-0.15	0.02	-0.07	0.03	-0.36^{*}	0.04	$-0.18^{\dagger}$	0.04
	Location × Verb Class	0.31	0.06	$0.31^{\dagger}$	0.05	-0.03	0.09	-0.19	0.1
500:600	Location	-0.25^{*}	0.08	-0.04	0.03	0.03	0.1	0.24^{*}	0.05
	Verb Class	-0.21	0.02	$-0.11^{\dagger}$	0.02	-0.43^{*}	0.06	$-0.17^{\dagger}$	0.04
	Location × Verb Class	0.36	0.07	$0.36^{\dagger}$	0.06	0.12	0.1	-0.09	0.1
600:700	Location	-0.25	0.08	-0.09	0.03	-0.06	0.1	0.33^{*}	0.05
	Verb Class	-0.16	0.02	$-0.12^{\dagger}$	0.02	-0.46^{*}	0.05	-0.21^{*}	0.04
	Location × Verb Class	0.16	0.07	$0.35^{\dagger}$	0.05	-0.17	0.08	0	0.11
700:800	Location	-0.18	0.08	-0.07	0.03	-0.03	0.1	0.32^{*}	0.05
	Verb Class	-0.08	0.02	$-0.11^{\dagger}$	0.02	-0.3^{*}	0.04	-0.14	0.04
	Location × Verb Class	0.17	0.06	0.22	0.05	0.04	0.07	-0.03	0.1
800:900	Location	-0.2	0.08	-0.06	0.02	-0.12	0.09	$0.17^{\dagger}$	0.04
	Verb Class	-0.1	0.02	-0.16^{**}	0.01	$-0.28^{\dagger}$	0.04	$-0.15^{\dagger}$	0.03
	Location × Verb Class	0.27	0.07	$0.28^{\dagger}$	0.05	0.18	0.06	0.24	0.08
900:1000	Location	-0.16	0.07	-0.02	0.02	-0.22	0.09	0.11	0.04

	Verb Class	-0.21*	0.02	-0.21**	0.02	-0.41*	0.02	-0.28**	0.03
	Location × Verb Class	0.54*	0.06	0.56**	0.05	0.13	0.04	0.3	0.08
1000:1100	Location	-0.19	0.07	0.02	0.03	-0.19	0.08	0.14	0.04
	Verb Class	-0.33**	0.02	-0.26**	0.02	-0.53**	0.03	-0.36**	0.03
	Location × Verb Class	0.81*	0.08	0.73**	0.05	0.25	0.07	0.35†	0.08
1100:1200	Location	-0.11	0.07	0.12	0.02	-0.05	0.08	0.19	0.05
	Verb Class	-0.27**	0.02	-0.23**	0.01	-0.57**	0.04	-0.38**	0.04
	Location × Verb Class	0.97**	0.08	0.8**	0.05	0.08	0.08	0.39	0.09
1200:1300	Location	-0.16	0.07	0.11	0.02	-0.01	0.08	0.26*	0.05
	Verb Class	-0.26**	0.01	-0.18**	0.01	-0.52**	0.03	-0.34**	0.03
	Location × Verb Class	0.71*	0.06	0.77**	0.05	-0.13	0.1	0.19	0.09
1300:1400	Location	-0.14	0.06	0.09	0.02	0.05	0.07	0.28*	0.04
	Verb Class	-0.23**	0.01	-0.17**	0.01	-0.34*	0.03	-0.25**	0.03
	Location × Verb Class	0.86**	0.06	0.89**	0.04	0.24	0.1	0.38	0.09
1400:1500	Location	-0.13	0.06	0.03	0.02	0.03	0.07	0.3*	0.05
	Verb Class	-0.17*	0.01	-0.14**	0.01	-0.28*	0.03	-0.23**	0.03
	Location × Verb Class	1.01**	0.06	0.96**	0.04	0.39	0.1	0.46†	0.09
1500:1600	Location	-0.19	0.06	-0.03	0.02	0.04	0.06	0.32*	0.05
	Verb Class	-0.22**	0.01	-0.14**	0.01	-0.21	0.03	-0.22**	0.03
	Location × Verb Class	1.1**	0.07	1.13**	0.04	0.54	0.1	0.62*	0.09
1600:1700	Location	-0.26	0.06	-0.07	0.02	0.15	0.06	0.34**	0.04
	Verb Class	-0.24**	0.01	-0.17**	0.01	-0.23*	0.02	-0.19**	0.02
	Location × Verb Class	1.37**	0.07	1.34**	0.04	0.7†	0.09	0.77**	0.08

1700:1800	Location	-0.28	0.06	-0.1	0.02	0.13	0.06	0.29*	0.04
	Verb Class	-0.22**	0.01	-0.15**	0.01	-0.22*	0.02	-0.21**	0.02
	Location × Verb Class	1.26**	0.08	1.35**	0.04	0.89*	0.1	0.99**	0.08
1800:1900	Location	-0.23	0.06	-0.03	0.02	0.19	0.06	0.35**	0.04
	Verb Class	-0.28**	0.02	-0.19**	0.01	-0.28**	0.02	-0.24**	0.02
	Location × Verb Class	1.38**	0.08	1.37**	0.05	0.96*	0.1	1.03**	0.08
1900:2000	Location	-0.23	0.06	-0.06	0.02	0.21	0.05	0.35**	0.04
	Verb Class	-0.2**	0.01	-0.19**	0.01	-0.29*	0.02	-0.25**	0.02
	Location × Verb Class	1.57**	0.08	1.55**	0.05	0.91*	0.1	0.97**	0.09
2000:2100	Location	-0.14	0.05	-0.04	0.03	0.23	0.06	0.35**	0.04
	Verb Class	-0.23**	0.02	-0.21**	0.02	-0.16	0.03	-0.16*	0.02
	Location × Verb Class	1.77**	0.09	1.72**	0.05	0.74†	0.1	0.85**	0.08
2100:2200	Location	-0.13	0.05	-0.17†	0.03	0.16	0.06	0.32*	0.04
	Verb Class	-0.19*	0.02	-0.3**	0.02	-0.19	0.03	-0.18**	0.02
	Location × Verb Class	1.78**	0.09	2.04**	0.06	0.8	0.11	0.89**	0.09
2200:2300	Location	-0.19	0.06	-0.3*	0.04	0.11	0.06	0.29*	0.05
	Verb Class	-0.13†	0.02	-0.3**	0.02	-0.25*	0.02	-0.18**	0.02
	Location × Verb Class	1.84**	0.08	2.34**	0.07	0.78	0.12	0.9**	0.09
2300:2400	Location	-0.25	0.06	-0.41**	0.04	0.16	0.06	0.39**	0.05
	Verb Class	-0.09	0.02	-0.27**	0.02	-0.26**	0.02	-0.2**	0.02
	Location × Verb Class	1.87**	0.09	2.61**	0.07	0.98*	0.12	0.93**	0.1
2400:2500	Location	-0.3	0.06	-0.38**	0.04	0.03	0.07	0.31*	0.05
	Verb Class	-0.09	0.02	-0.2**	0.02	-0.19*	0.02	-0.2**	0.02

	Location × Verb Class	2.05^{**}	0.09	2.71^{**}	0.08	1.05^{*}	0.12	1.02^{**}	0.1
2500:2600	Location	$-0.42^{\dagger}$	0.06	-0.34^{**}	0.04	-0.06	0.08	0.3^{*}	0.05
	Verb Class	-0.01	0.02	-0.1	0.02	-0.23^{**}	0.02	$-0.13^{\dagger}$	0.03
	Location × Verb Class	2.11^{**}	0.1	2.89^{**}	0.08	$0.99^{\dagger}$	0.12	0.96^{**}	0.11
2600:2700	Location	$-0.43^{\dagger}$	0.07	-0.04	0.05	-0.27	0.1	0.32^{*}	0.05
	Verb Class	0.01	0.02	-0.03	0.03	-0.29^{**}	0.03	-0.17^{*}	0.03
	Location × Verb Class	2.36^{**}	0.13	2.83^{**}	0.11	1.28^{*}	0.15	1.1^{**}	0.11
2700:2800	Location	$-0.46^{\dagger}$	0.09	0.31	0.07	-0.2	0.11	0.45^{**}	0.06
	Verb Class	-0.08	0.05	0.12	0.03	$-0.24^{\dagger}$	0.04	-0.15	0.04
	Location × Verb Class	2.39^{**}	0.18	2.71^{**}	0.16	1.58^{*}	0.2	1.69^{**}	0.14

Note. P-values were obtained by a model-comparison approach based on a likelihood-ratio $\chi^2(1)$ test of the change in the goodness of fit between the full model and a comparison model in which only the relevant fixed effect was removed

† $p < .10$

* $p < .05$

** $p < .01$